



Place de la structure génétique de l'espèce *Escherichia coli* dans l'état de son commensalisme intestinal et dans l'expression de sa virulence.

Mounira Smati

► To cite this version:

Mounira Smati. Place de la structure génétique de l'espèce *Escherichia coli* dans l'état de son commensalisme intestinal et dans l'expression de sa virulence.. Autre [q-bio.OT]. Université Paris-Nord - Paris XIII, 2014. Français. NNT : 2014PA132041 . tel-01231869

HAL Id: tel-01231869

<https://theses.hal.science/tel-01231869>

Submitted on 21 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE
pour obtenir le grade de
DOCTEUR en
Sciences de la Vie et de la Santé
École Doctorale : Galilée
Discipline : Biologie moléculaire

présentée et soutenue publiquement
par Mounira Smati
le 08 décembre 2014

Sujet :

**Place de la structure génétique de l'espèce *Escherichia coli*
dans l'état de son commensalisme intestinal et dans
l'expression de sa virulence**

Jury :

Pr Stéphane Bonacorsi :	Président
Pr Bertrand Picard :	Directeur de thèse
Pr Erick Denamur :	Co-directeur de thèse
Pr Michel Simonet:	Rapporteur
Pr Roland Quentin :	Rapporteur
Pr Olivier Schischmanoff :	Examineur

REMERCIEMENTS

Je remercie chaleureusement tous les membres du jury qui me font l'honneur d'examiner cette thèse :

- le Pr Stéphane Bonacorsi, qui a accepté de présider ce jury de thèse
- le Pr. Michel Simonet et le Pr. Roland Quentin, pour avoir accepté d'être rapporteurs de ce travail.
- le Pr. Olivier Schichmanoff

Je tiens à exprimer mes très sincères remerciements et ma plus grande reconnaissance à mes deux directeurs de thèse :

- Bertrand Picard,
- Erick Denamur.

Bien sûr, un grand merci à tous ceux avec qui j'ai travaillé en étroite collaboration pendant toutes ces années :

- Olivier Clermont, Sandrine Adiba....pour leur conseils et leur bonne humeur
- Odile Bouvet, merci pour ton soutien et ton écoute

Je remercie également Sara Dion, Frédéric Fourreau, Cécile Hignard, Etienne Girault, le personnel de l'ONF, le Pr. Sören Schubert, Giuseppe Magistro et le Pr. Pierre Cosson.

Merci aussi à Annie Vessières, sans qui ce travail n'aurait pas été possible.

A vous tous, amis et collègues de travail.

Un grand Merci.

TABLE DES MATIERES

INTRODUCTION GENERALE.....	8
PARTIE BIBLIOGRAPHIQUE	13
I. <i>E. coli</i> , un problème de santé publique	14
II. Diversité intra-spécifique du genre <i>Escherichia</i>	18
1). Le Genre <i>Escherichia</i>	18
a). Définition de l'espèce <i>E. coli</i>	18
b). Les espèces du genre <i>Escherichia</i>	19
2). Diversité intra <i>E. coli</i>	22
3). Diversité intra B2 et sous-groupes de B2	27
III. Commensalisme et <i>E. coli</i>	33
1). Distribution de <i>E. coli</i> dans la nature	33
2). Variabilité intra-hôte et inter-hôte	36
3). Diversité des phylogroupes	40
4). Aspects quantitatifs du commensalisme.....	46
IV. HPI et virulence	48
1). Influence du fond génétique dans l'expression des gènes de virulence .	48
2). Îlots de pathogénicité.....	51
3). Le PAI IV ou HPI.....	52
4). Structure du HPI	56
5). les autres sidérophores de <i>E. coli</i>	63
6). Virulence expérimentale liée au HPI	66
7). Modèle <i>Dictyostelium discoideum</i>	71
INTRODUCTION DE LA PARTIE EXPERIMENTALE	75
ARTICLE n°1	78

I. Introduction de l'article n°1	79
II. Article n°1	81
III. Commentaires de l'article n°1	90
ARTICLE n°2	102
I. Introduction de l'article n°2	103
II. Article n°2.....	105
III. Commentaires de l'article n°2	128
IMPACT DU HPI (ILOT DE HAUTE PATHOGENICITE OU PAIIV) SUR LA VIRULENCE	
<i>D'E. COLI</i> EN FONCTION DU FOND GENETIQUE:	
I. Introduction	136
II. Matériel et méthodes.....	139
1). Matériel	139
a). Souches bactériennes	139
b). Plasmides	146
c). Souches d'amibes	149
d). Souris	149
e). Milieux et antibiotiques	149
2). Méthodes :	150
a). Déterminants bactériens :	150
b). Vérification des résistances aux antibiotiques.....	153
c). Obtention des mutants et recombplémentation plasmidique .	153
d). Cure du plasmide pS88	154
e). Protocole de pathologie expérimentale avec <i>Dictyostelium discoideum</i>	154
f). Protocole de pathologie expérimentale avec la souris	157
g). Mesures du taux d'invasion intracellulaire	158
h). Analyses statistiques	159

III. Résultats	160
1). Détermination des facteurs de virulence	160
2). Virulence chez la souris dans un modèle de septicémie	162
a). Comparaison des souches sauvages	162
b). Souche 536 et ses mutants du HPI.....	163
c). Souche CFT073 et ses mutants du HPI	163
d). Souche NU14 et ses mutants du HPI	164
e). Souche C5 et ses mutants du HPI	168
f). Souche S88 et ses mutants du HPI.....	169
3). Résultats du modèle de virulence expérimentale chez l'amibe.....	171
a). Mutants du HPI de la souche de <i>E. coli</i> 536.....	171
b). Mutants du HPI de <i>E. coli</i> CFT073	171
c). Mutants du HPI de NU14	172
d). Mutants du HPI de <i>E. coli</i> C5.....	175
e). Mutants du HPI de <i>E. coli</i> S88	176
4). Mesures du taux d'invasion intracellulaire (intra-amibes)	177
IV. Discussion	181
CONCLUSION GENERALE	187
BIBLIOGRAPHIE	192

LISTE DES ABREVIATIONS

ACP : Analyse en Composantes Principales

ADN : Acide désoxyribonucléique

AFC : Analyse Factorielle des Correspondances

AIEC : *E. coli* adhérent et invasif

Ap : Ampicilline

BLSE : Bêta Lactamases à Spectre Elargi

DAEC : *E. coli* entéroadhérent

Cm : Chloramphénicol

Ct : cycle treshold

CV : coefficient de variation

DO : densité optique

EAEC : *E. coli* entéroaggrégatif

ECOR: « *E. coli* reference collection »

EHEC : *E. coli* entérohémorragique

EIEC : *E. coli* entéroinvasif

EPEC : *E. coli* entéropathogène

ETEC : *E. coli* entérotoxinogène

ExPEC : *E. coli* pathogènes extra intestinaux

FV : facteur de virulence

GR : Grazing résistant (résistant au broutage par l'amibe)

GS : Grazing sensible (sensible au broutage par l'amibe)

HDV : virus de l'hépatite Delta

HPI : « high pathogenicity island »

InPEC : *E. coli* pathogènes intestinaux

Kn : Kanamycine

LB : Lysogeny Broth

MCPB : Magnésium, Calcium, Phosphate Buffer

MLEE : Multilocus Enzyme Electrophoresis

MLST : Multilocus Sequence Typing

PAI : « pathogenicity islands » ou îlot de pathogénicité

PCR : Polymerase Chain Reaction

qPCR: PCR quantitative

RAPD: Randomly Amplified Polymorphic DNA

ROS: reactive oxygen species

rpm : rotation par minute

rrn RFLP : “restriction fragment length polymorphism des gènes des ARN ribosomiaux ou
ribotypage

ST : “Séquence type”

STEC : *E. coli* producteur de Shiga-toxine

Tet : Tétracycline

UFC : unités formant colonies

UPEC : *E. coli* uro-pathogènes

Zn : zéocine

INTRODUCTION GENERALE

Certaines espèces bactériennes présentent une spécificité d'hôte ou de comportement limitant leur écologie à des situations précises et stéréotypées. Ainsi *Neisseria gonorrhoeae* est un parasite strict de l'homme. D'autres sont, au contraire, ubiquitaires, saprophytes de l'environnement, commensales de l'homme et de nombreuses espèces animales, pathogènes dans un large répertoire de situations physiopathologiques. Parmi les pathogènes opportunistes, *Escherichia coli*, à la fois commensal et pathogène, est un modèle particulièrement intéressant par sa fréquence et sa responsabilité dans la survenue d'infections communautaires et nosocomiales intra et extra intestinales très diverses et par les connaissances accumulées sur cette bactérie modèle de la biologie, notamment en termes de phylogénie et de structure de population (Touchon *et al.*, 2009). Pour ces raisons, *E. coli* est le candidat idéal pour étudier la transition entre commensalisme et pathogénicité opportuniste (Pinheiro Da Silva *et al.*, 2007) ou encore comment le lien entre une bactérie et son hôte peut fluctuer entre mutualisme, commensalisme, pathogénèse opportuniste ou pathogénèse spécifique.

E. coli, est donc retrouvé comme véritable saprophyte de l'eau et des boues (Savageau *et al.*, 1983, Berthe *et al.*, 2013), comme commensal de l'homme et des animaux à sang chaud (Tenaillon *et al.*, 2010), comme pathogène intestinal et extra-intestinal (Russo and Johnson *et al.*, 2003).

Cette ubiquité peut s'expliquer par la diversité de l'espèce. Ainsi, un *E. coli* comporte de 4200 à 5500 gènes dans son génome et il existe dans l'espèce *E. coli* un total d'environ 20.000 gènes (Touchon *et al.*, 2009). Elle peut également traduire un intense pouvoir d'adaptation des clones qui la composent, aux différents compartiments où ils doivent survivre, se multiplier, qu'ils doivent coloniser. Ainsi, par exemple, dans le commensalisme intestinal, ces clones doivent inhiber leurs concurrents par leurs colicines, échapper aux prédateurs, phagocytes et amibes, par leurs enveloppes externes, utiliser au mieux les nutriments disponibles comme le gluconate du tube digestif (Chang *et al.*, 2004) .

Dans certaines circonstances de leur virulence, il leur faut entrer en concurrence avec les protéines de l'organisme qu'ils infectent pour capter des facteurs croissance essentiels comme le fer (Wandersman et Delepelaire *et al.*, 2004).

Cependant, les clones qui composent l'espèce se regroupent dans des structures plus larges, les groupes phylogénétiques qui sont maintenant au nombre de 8 (Moissenet *et al.*, 2010, Luo *et al.*, 2011, Clermont *et al.*, 2013) dont 4 principaux, A, B1, B2 et D (Selander *et al.*, 1986 ; Goullet et Picard 1989 ; Herzer *et al.*, 1990) et dont des études épidémiologiques ont montré qu'ils présentaient un certain degré de spécialisation en termes d'écologie et de pouvoir pathogène (Goullet et Picard 1986, Bidet *et al.*, 2007, Jauréguy *et al.*, 2008). Ainsi, les souches du groupe phylogénétiques B2 sont plus fréquemment impliquées dans les infections extra intestinales que celles des autres groupes (Escobar-Paramo *et al.*, 2004).

Plusieurs approches expérimentales sont utilisables pour étudier l'adaptation des clones de l'espèce *E. coli*. L'une, classique, qualitative, consiste à mettre en évidence des facteurs de colonisation ou de virulence, à étudier leurs effets dans les modèles cellulaires ou animaux, à détecter leur fréquence dans des collections d'isolats naturels. L'autre, maintenant possible grâce aux progrès de la biologie moléculaire, permet d'approcher la quantification de groupes bactériens dans un environnement donné. Ce qui était très limité avec les techniques de dénombrement phénotypique lourdes et applicables de faibles échantillons peut être élargi par les techniques de PCR en temps réel. Cela conduit à analyser les équilibres entre les populations à l'intérieur d'une espèce, entre les espèces, et à définir en situation commensale des profils d'abondance d'espèces caractéristiques d'une niche écologique donnée.

L'effet de l'inoculum infectant est peu documenté. Il a été étudié dans certaines situations comme les infections urinaires et les infections survenant sur cathéters où la notion de seuils significatifs est maintenant établie. Mais son importance reste mal connue en physiologie digestive et inconnue dans de nombreuses situations physiopathologiques (portes d'entrée des

bactériémies, abcès intra abdominaux, infections des séreuses par exemple) et est encore à approfondir.

Son étude pourrait permettre de définir des seuils quantitatifs critiques lors du franchissement d'une barrière anatomique ou physiologique ou lors de l'aggravation d'une situation clinique.

Beaucoup d'études ont été menées sur les souches de *E. coli* pathogènes mais peu ont été consacrées aux souches commensales. Il apparaît pourtant important d'élucider la structure des populations commensales tant du point de vue écologique qu'évolutif afin de mieux comprendre la virulence et la résistance aux antibiotiques. En effet, la pression de sélection au niveau de la flore peut promouvoir l'émergence de facteurs de virulence et /ou de résistance aux antibiotiques (notion de réservoir) (Skurnik *et al.*, 2009; Fantin *et al.*, 2009).

L'objectif de cette thèse est d'utiliser les 2 approches, quantitative et qualitative, pour l'étude de l'adaptation de *E. coli* et les rapports de cette adaptation avec la structure génétique de l'espèce dans deux circonstances particulièrement caractéristiques, le commensalisme intestinal et la virulence extra-intestinale. Après une présentation des données de la littérature concernant la structure de population de *E. coli*, les données de son commensalisme et la place du métabolisme du fer dans le commensalisme et la virulence de l'espèce, nous avons développé deux types d'expérimentations.

Nous avons mis au point une technique de PCR en temps réel, permettant de quantifier les 4 principaux groupes phylogénétiques de l'espèce dans les selles de plusieurs hôtes de *E. coli* : l'homme et des espèces d'animaux à sang chaud, choisies pour la diversité de leur taille, de leur régime et de leur mode de vie afin de caractériser des structures de populations commensales définies par les proportions variables des 4 groupes phylogénétiques.

Pour étudier la virulence, nous avons choisi l'un des domaines les plus critiques qui conditionnent l'adaptation d'un clone virulent dans un compartiment extra-intestinal, le métabolisme du fer.

En effet, il existe chez *E. coli* 4 systèmes de captation du fer souvent simultanément présents dans les souches (Braun *et al.*, 2002), et des travaux récents portant sur la comparaison de très larges collections de souches commensales et de souches virulentes montrent que parmi une vingtaine de facteurs de virulence de diverses natures, les systèmes de captation du fer étaient les plus différenciant (A.S Daubié et E. Denamur, communication personnelle). Parmi ces systèmes, nous avons étudié la yersiniabactine, un sidérophore dont la synthèse, le transport et la régulation sont codés par des gènes situés dans le « High Pathogenicity Island » (HPI) dont l'expression présente l'avantage d'être critique cliniquement chez l'homme (Clermont *et al.*, 2001, Jonhson *et al.*, 2001) et dans différents modèles de virulence expérimentale, sur la souris (Schubert *et al.*, 2002) et sur l'amibe *Dictyostelium discoideum* (Adiba *et al.*, 2010). Nous avons analysé les rapports du HPI et de son expression dans ces 2 modèles avec le fond génétique de clones caractérisé par leur appartenance à différents sous-groupes du groupe phylogénétique B2.

PARTIE BIBLIOGRAPHIQUE

I. *E. coli*, un problème de santé publique

Les bactéries de l'espèce *E. coli* posent un réel problème de santé publique à l'échelon mondial. Chaque année, *E. coli* cause plus de deux millions de décès dus aux diarrhées infantiles et aux infections extra intestinales (essentiellement les septicémies ayant pour point de départ une infection urinaire) (Kosek *et al.*, 2003; Russo et Johnson 2003).

E. coli se comporte comme un pathogène opportuniste chez les humains et chez les animaux domestiques, mais aussi comme un pathogène obligatoire spécifique de l'homme (*E. coli* entéro-invasif ETEC et *Shigella*) (Orskov et Orskov 1992).

Shigella, qui a été longtemps considérée comme un genre avec ses quatre espèces (*dysenteriae*, *flexneri*, *boydii*, and *sonnei*) basée sur leur capacité à générer une diarrhée invasive et sur des caractères biochimiques propres, appartient en fait à l'espèce *Escherichia coli*. La dysenterie bacillaire est une maladie diarrhéique sévère due à *Shigella* affectant des centaines de millions de personnes de par le monde et causant plus de 500.000 décès par an (Berche *et al.*, 1988). La souche entéro-hémorragique O157:H7 a récemment été l'agent d'épisodes de diarrhées sanglantes et de syndromes hémolytiques et urémiques lors d'intoxications alimentaires aux Etats Unis, en Europe et en Asie.

Les infections extra-intestinales sont aussi très variées. Les plus fréquentes sont les infections urinaires, les septicémies et les méningites. *E. coli* est responsable de plus de 80 % des 150 millions de cystites simples survenant chaque année dans le monde (Stamm et Hooton *et al.*, 1993 ; Russo et Johnson 2003). Chaque épisode entraîne en moyenne six jours de symptômes et deux jours d'activité réduite au travail (Foxman et Frerichs 1985), dont le coût économique et social n'a pas été évalué. C'est aussi l'agent pathogène retrouvé dans 90 % des pyélonéphrites (Talan *et al.* 2000). Il serait donc responsable de 5,4 millions de pyélonéphrites chaque année (Russo et Johnson 2003). Là encore, l'arsenal moléculaire requis par les souches uropathogènes est impressionnant et a fait l'objet d'une dissection moléculaire méticuleuse (Bahrani-Mougeot *et al.* 2002 ; Wiles *et al.* 2008).

Bien qu'aucun chiffre directement mesuré ne soit disponible, des estimations indiquent que plus de 800 000 personnes mourraient de septicémie à *E. coli* chaque année dans le monde (Russo et Johnson 2003). Ce chiffre à lui seul ferait du sepsis sévère à *E. coli* la septième cause de mortalité infectieuse dans le monde, après les infections respiratoires basses (4 millions de morts par an), le VIH (2,6 millions de morts), les diarrhées (2,1 millions de morts, dont un nombre substantiel sont dues à *E. coli* et à *Shigella*), la tuberculose (1,7 millions de morts), le paludisme (un million de mort) et la rougeole (875 milles morts) (WHO 2000). Si l'on combine les décès secondaires aux pathologies intestinales et extra-intestinales, *E. coli* est l'une des bactéries les plus mortelles au monde avec l'agent de la tuberculose.

Enfin, les méningites néonatales bien que rares (400 cas par an aux États-Unis d'Amérique chaque année) (Russo et Johnson 2003), sont d'un pronostic effroyable. Plus d'un tiers des nouveau-nés atteints meurent et la moitié des survivants développent d'importantes séquelles (De Louvois 1994). Environ 40 % sont dus à *E. coli*. Cette proportion ne cesse d'augmenter du fait de l'utilisation fréquente de la pénicilline A chez les femmes enceintes porteuses de Streptocoques du groupe B (agent principal des méningites néonatales) (Stoll *et al.* 2002).

L'étendue des infections extra-intestinales dues à *E. coli* est encore vaste et comporte des pneumopathies, des infections-abdominopelviennes, des infections des parties molles, des ostéomyélites, des infections du site chirurgical. Le développement de procédures thérapeutiques toujours plus invasives a aussi fait d'*E. coli* un agent non négligeable d'infections nosocomiales et iatrogènes : infections sur cathéter central, sur sonde urinaire, pneumonie acquise sous ventilation mécanique (Russo et Johnson 2003).

Le contrôle de l'antibiorésistance est un défi majeur en matière de santé publique. Hors, durant la dernière décennie, on a assisté chez les entérobactéries à une véritable pandémie des bêtalactamases à spectre élargi (BLSE) de type CTX-M conduisant à des impasses

thérapeutiques (Livermore et Woodford 2006). *E. coli* est l'entérobactérie la plus affectée par ce type de BLSE, qui touche aussi bien les souches pathogènes que commensales.

Récemment, des clones épidémiques majeurs de *E. coli* produisant une BLSE de type CTX-M émergent, tel que le clone virulent O25b: H4-ST131 produisant une BLSE de type CTX-M15, clone de répartition mondiale, (Nicolas-Chanoine *et al.*, 2008; Clermont *et al.*, 2008; Vimont *et al.*, 2012) et le clone O102:H6- ST964 (Deschamps *et al.*, 2009) sont apparus.

Les souches pathogènes ont été classées en différents pathovars basé sur les conditions d'isolement et la présence de facteurs de virulence spécifiques (Croxen et Finlay 2010).

E. coli est donc l'une des espèces les plus fréquemment rencontrées en pathologie humaine intestinale et extra-intestinale (Kaper 2004).

On distingue les souches de *E. coli* pathogènes selon la localisation intestinale ou extra-intestinale des infections qu'elles produisent.

Les souches de *E. coli* pathogènes extra-intestinales sont regroupées sous la dénomination commune ExPEC (Extra intestinal Pathogenic *E. coli*) (Russo and Johnson 2000).

On distingue essentiellement deux pathovars au sein des ExPEC :

- les *E. coli* uro-pathogènes (UPEC),
- les *E. coli* responsables de méningites, le plus souvent chez le nouveau-né.

Les infections extra-intestinales à *E. coli* sont communes à tous les groupes d'âge et peuvent mettre en jeu quasiment n'importe quel organe ou site anatomique chez l'homme. Les infections extra-intestinales à *E. coli* regroupent des infections du tractus urinaire, des méningites (le plus souvent chez le nouveau-né), diverses infections intra-abdominales, des pneumonies (chez les patients hospitalisés), des infections sur dispositifs intravasculaires, des ostéomyélites et des infections des tissus mous. Des bactériémies peuvent compliquer l'infection de l'ensemble de ces sites. Les ExPEC sont incapables de produire des infections intestinales, mais peuvent coloniser le tractus intestinal.

Par analogie avec la terminologie des souches pathogènes extra-intestinales ExPEC, les souches de *E. coli* pathogènes intestinales sont regroupées sous la dénomination commune InPEC (Intestinal Pathogenic *E. coli*). On distingue six pathovars au sein des InPEC :

- les *E. coli* entéro-toxinogènes (ETEC),
- les *E. coli* entéro-hémorragiques (EHEC),
- les *E. coli* entéro-invasifs (EIEC) et les *Shigella*,
- les *E. coli* entéro-pathogènes (EPEC),
- les *E. coli* à adhérence diffuse (DAEC),
- les *E. coli* adhérents et invasifs (AIEC) (Darfeuille-Michaud, 2002)
- et les *E. coli* entéro-agrégatifs (EAEC) (Nataro et Kaper 1998).

Les souches EIEC, *Shigella* sont des pathogènes obligatoires, c'est à dire qu'elles ne sont quasiment jamais trouvées à l'état commensal dans la flore intestinale. A l'inverse, les souches EAEC, DAEC, EPEC ainsi que les ExPEC sont des pathogènes non obligatoires c'est à dire qu'elles peuvent être retrouvées dans les selles des sujets sains avec une fréquence variable selon les individus et les populations humaines étudiées (Nataro et Kaper 1998).

II. Diversité intra-spécifique du genre *Escherichia*

L'espèce est, en bactériologie, l'unité de base de toute démarche taxonomique.

Il convient de définir l'espèce *E. coli* dans le phylum supérieur qui l'inclut, le genre *Escherichia*. puis de voir l'apport de la biologie moléculaire à la caractérisation intra spécifique ; les entités à décrire devenant les groupes phylogénétiques et plus récemment à l'intérieur du groupe B2, les sous-groupes.

Dans cette espèce très polymorphe, ces sous-groupes sont encore de vastes ensembles de clones dans lesquels des techniques plus discriminantes comme le MultiLocus Sequence Typing (MLST) distinguent de nombreuses séquences-types (ST). Le séquençage nucléotidique complet du génome permettra la différenciation ultime des souches.

1). Le Genre *Escherichia*

a). Définition de l'espèce *E. coli*

Il faut tout d'abord rappeler les principes qui ont permis la définition des espèces. Historiquement, celles-ci ont pu être déterminées à partir de caractères phénotypiques. Des individus d'une même espèce présentent des caractères morphologiques communs dans des zones géographiques particulières et distinctes d'autres individus.

En microbiologie, les caractères biochimiques et physiologiques ont été associés aux caractères morphologiques pour définir les espèces (Sneath et Sokal 1973). Pour *E. coli*, un bacille à Gram négatif aéro-anaérobie facultatif, ces derniers comprennent la production d'indole, la présence de différents enzymes telles que la lysine décarboxylase, la β -xylosidase, la β -glucuronidase, l'utilisation du sorbitol, du malonate et de l'adonitol (Farmer *et al.*, 1985). La prise en compte de cette définition impliquait alors une différenciation des *Shigella* présentant des caractères morphologiques, biochimiques et de pathogénicité particuliers.

Par la suite, les espèces ont pu être caractérisées selon la définition phénétique, proposée par Brenner et Falkow (Brenner *et al.*, 1972, Grimont *et al.*, 1988). Cette définition indique qu'un groupe de souches appartiennent à la même espèce si plus de 70 % de leurs ADN s'hybrident

ensemble dans des conditions peu restrictives et la stabilité thermique des brins hybridés réassociés n'est pas différente de plus de 5°C par rapport à des brins réassociés homologues. Ces observations ont pu être confirmées par la suite à l'aide du séquençage des gènes puis des génomes complets de souches de *E. coli* (Touchon *et al.*, 2009).

Les hybridations ADN/ADN rendant indistinguables les *Shigella* des souches de *E. coli* ; leur classification à l'intérieur de cette espèce est donc incontestable (Brenner *et al.*, 1972). Par la suite, les données obtenues par différents marqueurs moléculaires : multi locus enzyme electrophoresis (MLEE) (Branger *et al.*, 2005), ribotypage (Desjardins *et al.*, 1995), multilocus sequence typing (MLST) (Escobar-Paramo *et al.*, 2004a), et la comparaison de génomes entiers ont pu confirmer ces observations.

b). Les espèces du genre *Escherichia*:

E. coli a été décrit comme la première espèce du genre *Escherichia* par Castellani et Chambers en 1919.

En 1973, Burgess *et al.* ont décrit deux biotypes isolés de cafards et ont proposé une nouvelle espèce dénommée *Escherichia blattae* (Burgess 1973). Cette désignation était basée sur une similarité phénotypique avec *E. coli*. Depuis, cette espèce a été reclassée et est nommée *Shimwellia blattae* (Priest et Barker, 2010). En 1982, l'hybridation ADN-ADN a été utilisée comme un second critère dans la description de *E. hermanni* (Brenner *et al.*, 1982) et de *E. vulneris* (Brenner *et al.*, 1982). Basée sur un faible nombre d'isolats, ces deux espèces partagent un nombre minimal de similarités génétiques avec *E. coli* (~39 % de similarité) comparé aux autres membres de la famille des entérobactéries. Actuellement, les données phylogénétiques démontrent qu'ils n'appartiennent pas au genre *Escherichia* (Pham *et al.*, 2007, Paradis *et al.*, 2005). Pour les deux nouvelles espèces, *E. fergusonii* (Farmer *et al.*, 1985) et *E. albertii* (Huys *et al.*, 2003), les données phénotypiques, l'hybridation ADN, et les données phylogénétiques confirment l'appartenance au genre *Escherichia*. *E. fergusonii* est un pathogène opportuniste extraintestinal chez l'homme (Farmer *et al.*, 1985, Funke *et al.*, 1993, Savini *et al.*, 2008), et chez les animaux (Herraez *et al.*, 2005, Hariharan *et al.*, 2007). *E. albertii* est responsable de diarrhées

chez l'enfant (Huys *et al.*, 2003). Les hybridations montrent un grand degré de similarité génétique à *E. coli* (~60 % de similarité), et les données phylogénétiques regroupent *E. fergusonii* avec *E. coli* (Lawrence *et al.*, 1991). L'analyse des mêmes loci n'a pas été effectuée pour *E. albertii*, mais Hyma *et al.* ont utilisé la méthode MLST pour démontrer que les souches représentatives étaient monophylétiques et partageaient un ancêtre commun avec *E. coli* après la séparation *Salmonella-Escherichia* (Hyma *et al.*, 2005).

Plus récemment en 2009, cinq nouveaux clades de *E. coli* nommés CI à CV ont été identifiés et caractérisés génétiquement (Walk *et al.*, 2009). Ces souches sont phénotypiquement caractérisées comme des *E. coli* (lactose +, citrate -, indole +) mais par analyses MLST, il a été démontré que ces clades forment des groupes génétiques bien distincts (Figure 1). Ces clades accroissent la diversité de l'espèce. Depuis 2011, l'hypothèse que les souches de clade I constituent le huitième groupe phylogénétique de *E. coli* a été avancée sur la base d'un haut niveau de recombinaison entre ce clade I et *E. coli* (Luo *et al.*, 2011).

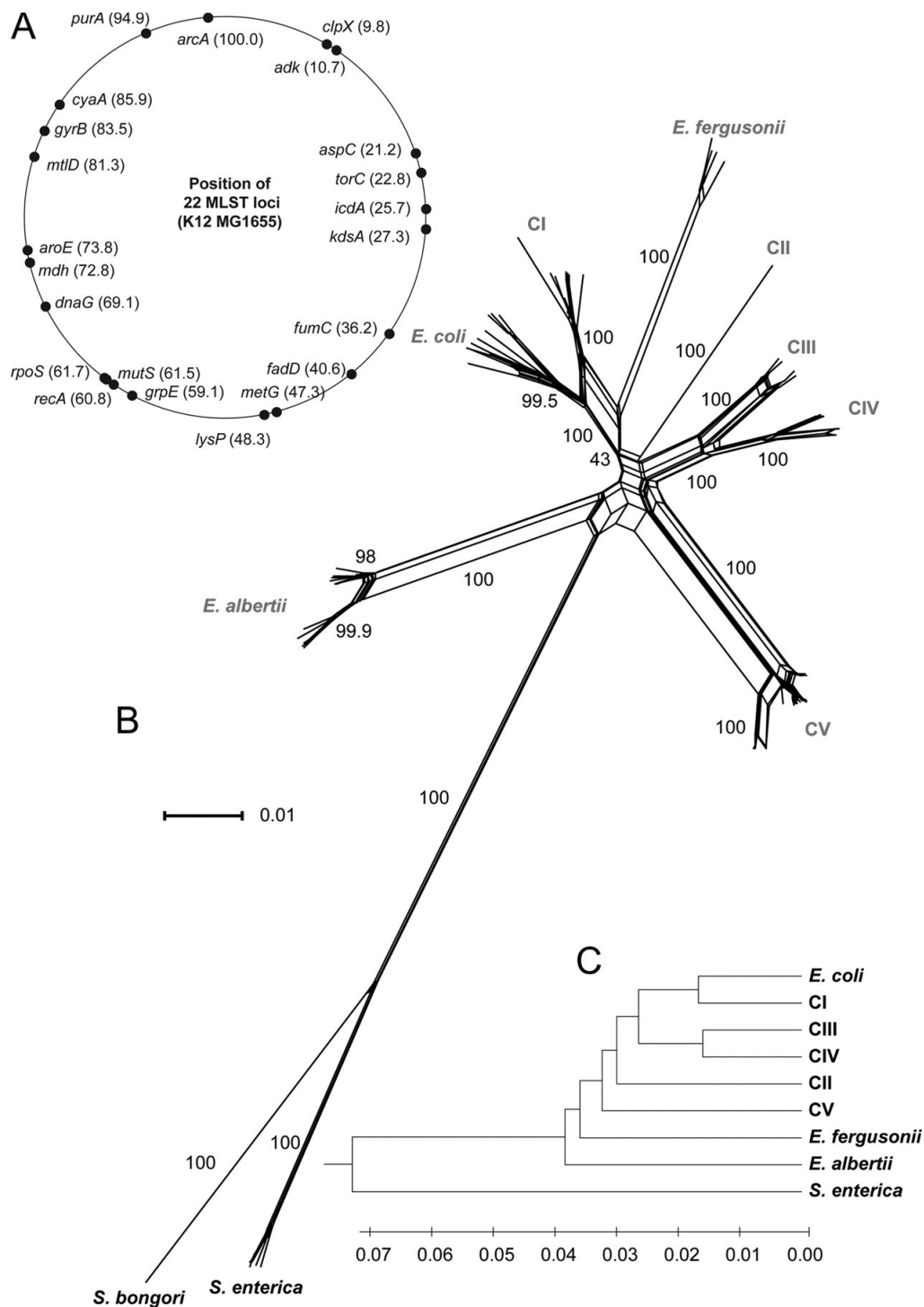


Figure 1 : Relations phylogénétiques des nouveaux clades de *Escherichia*

(A) Position des gènes de ménage utilisés dans les schémas de MLST dans le génome de *E. coli* K-12 MG1655. (B) Dendrogramme des isolats indiquant la position phylogénétique des 5 nouveaux clades par rapport aux autres espèces du genre *Escherichia*. (C) Arbre phylogénétique correspondant (d'après Walk *et al.* 2009).

2). Diversité intra *E. coli*

E. coli est l'une des espèces bactériennes parmi les plus étudiées et les mieux connues (Neidhardt *et al.*, 1996). La séquence complète du génome de nombreux représentants de cette espèce est disponible (plusieurs centaines de génomes de *E. coli* actuellement).

Dès le début des années 1980 avec les travaux de Thomas Whittam, Howard Ochman et Robert Selander (Whittam *et al.*, 1983) l'existence d'une structure génétique au sein de *Escherichia coli* a été révélée. Les dernières études confirment l'existence d'une sous-structure extensive de l'espèce (revue faite par Chaudhuri et Henderson en 2012).

La structure de population de l'espèce *E. coli* / *Shigella* est de type essentiellement clonale avec toutefois l'existence de transferts horizontaux (Desjardins *et al.*, 1995). Plusieurs collections ont été réunies pour représenter l'espèce, la plus utilisée est la collection ECOR (*E. coli* reference) représentative de la diversité génétique de l'espèce et pour lesquelles de nombreuses données génétiques et phénotypiques sont disponibles. Cette collection établie par Ochman et Selander en 1984 (Ochman *et al.*, 1984), comporte 72 souches sélectionnées à partir d'une collection précédemment établie de 2600 isolats naturels dont beaucoup avaient été fournis par Milkman (Milkman *et al.*, 1973). Ces 72 souches proviennent de différentes zones géographiques et se répartissent en 61 isolées de la flore fécale de sujets sains (29 humaines et 32 issues de 16 espèces animales différentes) et 11 isolées en Suède au cours de cystites ou de pyélonéphrites aiguës chez des humains.

La sélection de ces souches a été basée sur leur variabilité électrophorétique pour 11 enzymes métaboliques (MLEE) et pour leur présence chez des hôtes et à des sites géographiques les plus divers possibles; l'objectif étant d'établir une collection de souches représentatives de la diversité génétique de l'espèce.

La liste des souches ECOR ainsi que quelques-unes de leurs caractéristiques respectives sont disponibles à l'adresse internet <http://www.foodsafe.msu.edu/whittam/ecor/>.

Les premiers phénogrammes obtenus grâce aux données du MLEE identifiaient 4 groupes principaux (A, B1, B2 et D) au sein de la collection ECOR (Selander *et al.*, 1986 ; Goullet et Picard 1989 ; Herzer *et al.*, 1990). Puis, les auteurs se sont intéressés à la concaténation de différents gènes du multi-locus sequence typing (MLST). Les mêmes groupes ont été retrouvés que ce soit par des approches phylogénétiques ou de génétique des populations (Wirth *et al.*, 2006 ; Lecointre *et al.*, 1998 ; Escobar-Paramo *et al.*, 2004 ; Johnson *et al.*, 2006). La phylogénie de *E. coli* se précise alors, également à l'aide du MLST sur un nombre de souches plus important (Reid *et al.*, 2000 ; Hershberg *et al.*, 2007) et en utilisant *E. fergusonii* comme groupe externe (Escobar-Paramo *et al.*, 2004b). L'existence du phylogroupe E, dont font partie certaines souches 0157 :H7 est à présent établie (Tenaillon *et al.*, 2010). Plus récemment, le phylogroupe C a été proposé pour un groupe de souches proches mais distinctes du phylogroupe B1 (Moissenet *et al.*, 2010, Clermont *et al.*, 2011).

A la base de cet arbre, le groupe des B2 apparaît le plus diversifié avec au moins 9 sous-groupe phylogénétiques observés (Le Gall *et al.*, 2007). Puis, un sous-groupe du groupe D (appelé F) se distingue (Jaureguay *et al.*, 2008), ce phylogroupe est à présent bien documenté et forme un groupe frère avec B2 (Clermont *et al.*, 2011). Le reste du groupe D émerge d'abord, suivi par le groupe E. Finalement, les groupes frères A et B1 apparaissent (Figures 2 et 3).

La phylogénie de l'espèce a été confirmée par la comparaison de 20 génomes entiers (Touchon *et al.*, 2009).

On considère à l'heure actuelle que cette espèce bactérienne comprend 8 groupes phylogénétiques principaux avec sept groupes A, B1, B2, C, D, E et F appartenant à *E. coli sensu stricto* (Moissenet *et al.*, 2010) et un correspondant à *E. coli* clade I (Luo *et al.*, 2011, Clermont *et al.*, 2013).

La technique du MLST permet d'affilier les souches à l'un de ces 7 groupes phylogénétiques, ainsi que la dernière méthode Clermont (Clermont *et al.*, 2013) basée sur l'amplification ou non des gènes *arpA*, *chuA*, *yjaA* et du fragment d'ADN TspE4.C2. Cette

méthode permet aussi d'identifier des isolats membres des clades cryptiques II à V. Elle assigne correctement 95% des isolats de *E. coli* en comparaison au MLST.

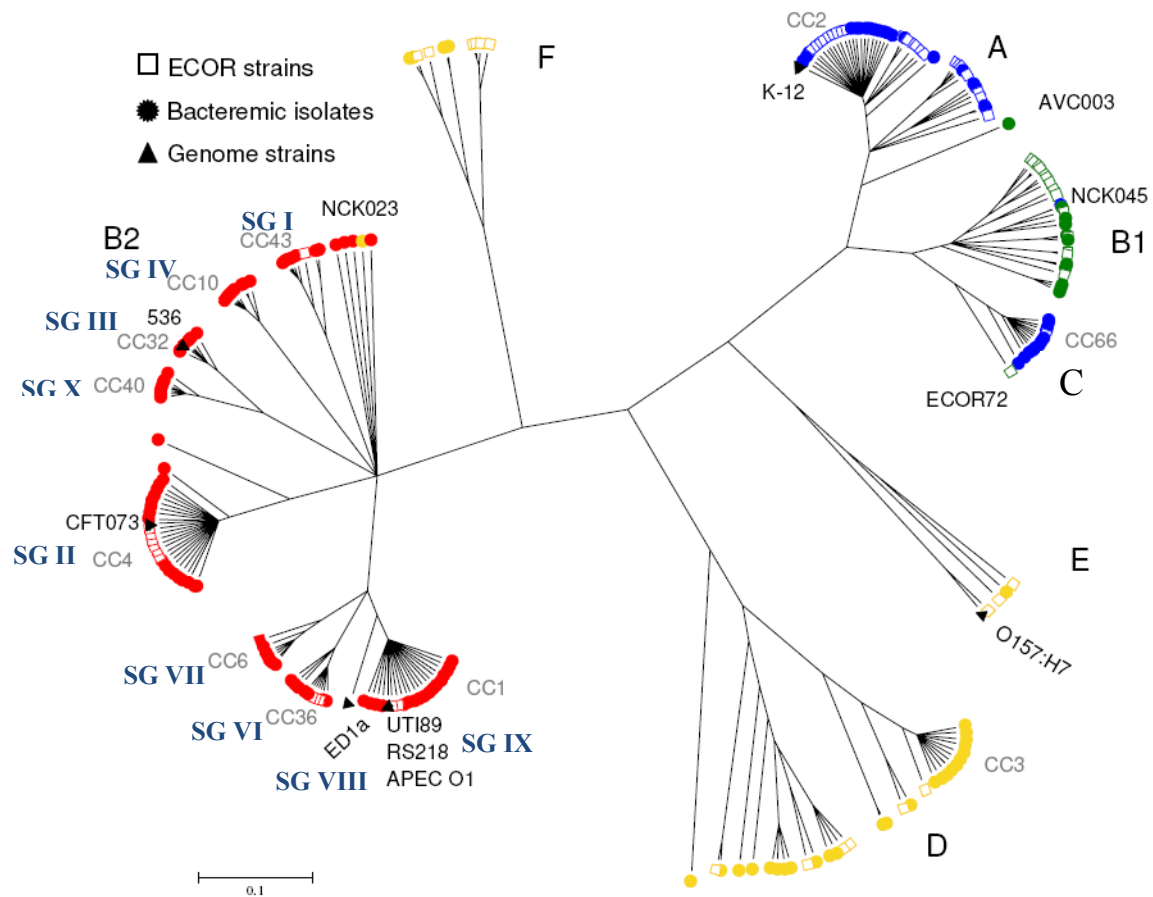


Figure 2: La phylogénie de *E. coli*. Analyse phylogénétique réalisée avec Clonalframe basée sur les séquences de 8 gènes de ménages de 161 souches de *E. coli* issus de bactériémie (cercles) et 67 souches de la collection de référence ECOR (carrés), ainsi que de 7 génomes de références (triangles). En bleu, sont indiqués les sous-groupes (SG) du phylogroupe B2 (adapté de Jauréguy *et al.*, 2008).

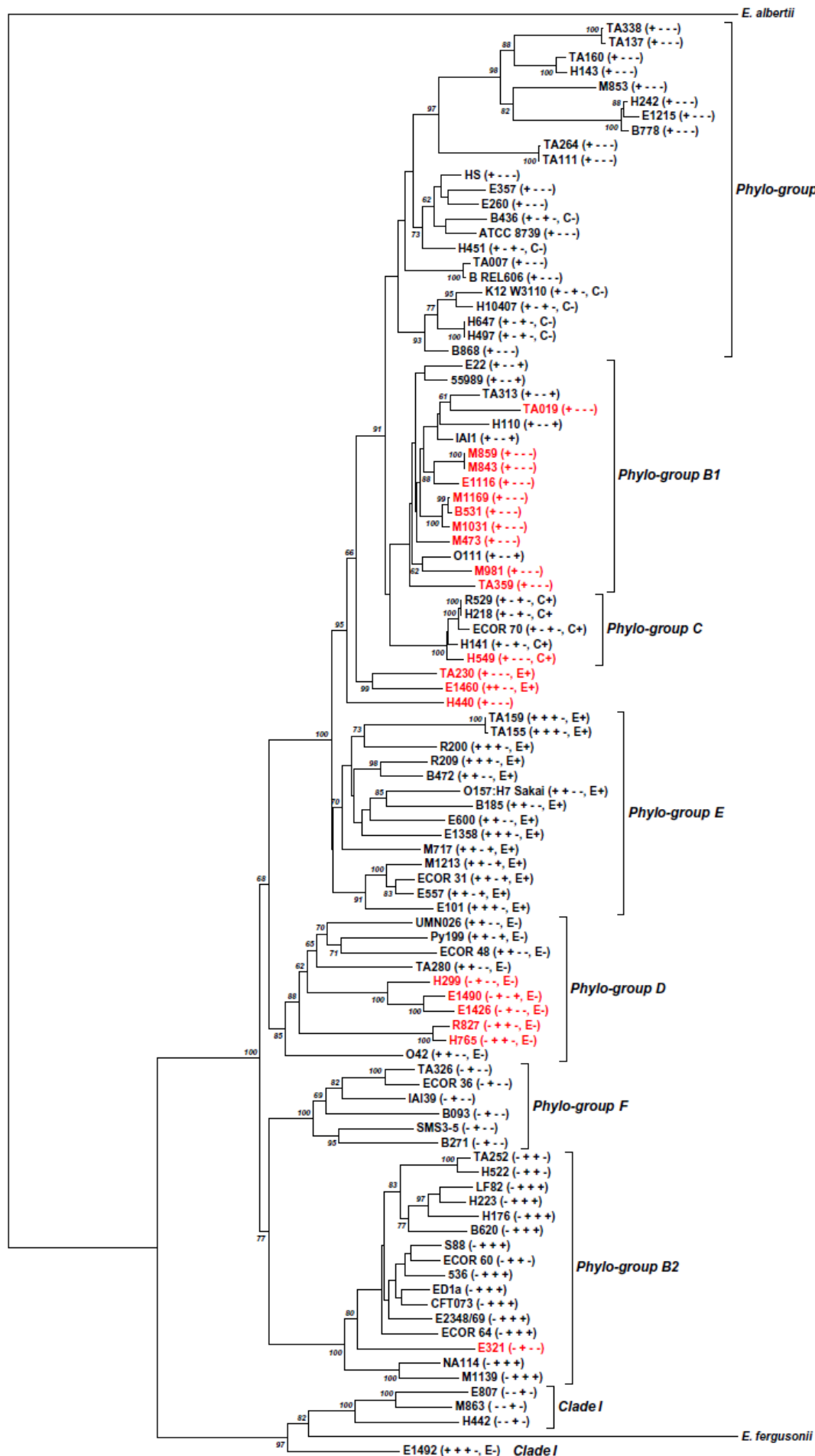


Figure 3 : Arbre phylogénétique de souches de *E.coli* représentatives de l'espèce et d'une souche de *E. fergusonii* (Clermont *et al.*, 2013).

L'arbre a été construit par la méthode de parcimonie à partir de des séquences de 13 gènes chromosomiques *trpA*, *trpB*, *pabB*, *putP*, *icd*, *polB* (Le Gall *et al.*, 2007), *aes* (Lescat 2009), *adk*, *fumC*, *gyrB*, *mdh*, *purA* et *recA* qui sont utilisés dans le schéma MLST de l' University College Cork (<http://mlst.ucc.ie/mlst/dbs/Ecoli>). La souche de *E. albertii* a servi pour enraciner l'arbre. Les groupes phylogénétiques A, B1, B2, C, D, E, F et le clade I sont indiqués. Les souches indiquées en rouge correspondent à des souches mal assignées par la nouvelle méthode Clermont quadruplex.

3). Diversité intra B2 et sous-groupes de B2 :

Les souches du groupe phylogénétique B2 et dans une moindre mesure du groupe D constituent au sein de l'espèce *E. coli* deux groupes clairement séparés des autres à la fois en terme de phylogénie mais également en terme de pathogénicité extraintestinale.

Une première analyse génétique de 60 souches du groupe phylogénétique B2 sélectionnées pour leur diversité écologique et génétique a été réalisée par MLST (Le Gall *et al.* 2007).

Cette étude a permis la délimitation de 9 sous-groupes notés I à IX (Figure 4).

Toutes les souches provenaient de localisations géographiques diverses. La plupart (49/60) étaient regroupées en sous-groupes présentant des caractéristiques variées :

- Les souches du sous-groupe I : d'origine humaine et animale, isolées dans des situations commensales et d'infections urinaires. C'est à ce sous-groupe qu'appartient le clone ST131-O25b, de répartition mondiale et porteur de la bêta-lactamase à spectre étendu (BLSE) CTX-M15 (Vimont *et al.*, 2012), mais y appartiennent aussi des isolats sans signification clinique. Les types O sont O6, O16, O25b (dans 40% des cas). Il correspond au STc131 dans le schéma d'Achtman (www.mlst.net).
- Les souches du sous-groupe II : d'origine principalement humaine mais aussi animale, principalement commensales, mais aussi isolées au cours d'infection urinaires ou de septicémie. Les types O sont O2a, O6, O25a (dans 100 % des cas). Il correspond au STc73.
- Les souches du sous-groupe III : d'origine principalement humaine mais également animale et avaient été isolées comme commensales, agents de septicémie ou d'infection urinaire, elles sont à 100% du type O6. Il correspond au STc127 dans le schéma d'Achtman.
- Les souches du sous-groupe IV : d'origine principalement humaine mais aussi animale et isolées lors d'infections ou en situation commensale, majoritairement O2b (dans 40% des cas). Il correspond au STc141.

- Les souches du sous-groupe V : d'origine était strictement humaine, et provenaient de situations commensales et d'infections d'urinaires, à 100% O16. Il correspond au STc144.
- Les souches du sous-groupe VI : d'origine strictement humaine et avaient été prélevées dans des situations commensales et lors d'infections urinaires, majoritairement O4 (85,7% des cas). Il correspond au STc12.
- Les souches du sous-groupe VII : d'origine strictement humaine et provenaient de situations pathogènes, urinaires ou intestinales, O75 dans 60% des cas. Il correspond au STc14.
- Les souches du sous-groupe VIII : d'origine strictement humaine et avaient été isolées de situations commensales digestives, toutes O81 (Clermont *et al.*, 2008). Il correspond au STc452.
- Les souches du sous-groupe IX : d'origine strictement humaine, et avaient été isolées de situations commensales, d'infections urinaires, de septicémies et surtout de méningites néo-natales. De types O majoritairement O1, O2a, O18 et O45a. Il correspond au STc95 dans le schéma d'Achtman.

Par ailleurs, des travaux de caractérisation de ces souches dans un modèle murin de septicémie, de contenu en virulence intrinsèque et de sérogroupage ont amené les observations suivantes. Les souches des sous-groupes I, III et IV comportaient des souches strictement virulentes, alors que les sous-groupes II, V, VI, VII et IX comportaient des souches virulentes et avirulentes chez la souris. Cette virulence a été globalement retrouvée chez tous les sous-groupes hormis le sous-groupe VIII, comprenant uniquement des souches commensales avirulentes chez la souris et auquel appartient la souche ED1a. Enfin, il a été remarqué que certains sous-groupes présentaient des sérogroupes particuliers. Ainsi les sous-groupes III, VI, VII et VIII présentaient les types O O6, O4, O75 et O81, respectivement.

Depuis, grâce à l'ajout d'analyses épidémiologiques par MLST, il apparaît que les sous-groupes II et IX (STc73 et STc95 comme définies dans le schéma de MLST d'Achtman) sont

trouvées respectivement à haute fréquence dans les méningites néonatales (Bidet *et al.*, 2007), dans les bactériémies (Bert *et al.*, 2010, Mahjoub-Messai *et al.*, 2011) et les infections du tractus urinaire (Gibreel *et al.*, 2012). On note en particulier l'association du sous-groupe IX avec les méningites néonatales. Ces deux sous-groupes représentent près de la moitié des infections extra-intestinales (Clermont *et al.*, 2014).

Plus récemment, un dixième sous-groupe a été décrit (sous-groupe X) comprenant majoritairement les souches O83, et représenté par la souche archétypale A0 34/86 (Clermont *et al.*, 2014) auquel s'ajoute le sous-groupe EPEC-1 (entéropathogène).

Il est également intéressant de noter que ces sous-groupes ont une diversité intraspécifique variable, allant de la quasi-clonalité avec très peu de STs (sous-groupes II, III, VIII, IX), à une diversité plus grande pour les sous-groupes I et VII.

Il a donc été démontré, comme pour l'espèce entière, que malgré des transferts horizontaux de gènes assez fréquents, le groupe phylogénétique B2 avait une structure clonale claire. Les 10 sous-groupes présentent une grande diversité génétique dans leur génome variable. La divergence nucléotidique sur gènes du core génome (partie du génome présente dans toutes les souches) entre deux souches B2 peut atteindre 2 %, contre 3 % pour deux souches dans toute l'espèce.

Sur la plan de la virulence, les trois-quarts des souches du groupe B2 sont virulentes lorsqu'elles sont testées dans le modèle de septicémie de la souris (décrit par Picard *et al.*, 1999). Le sous-groupe I, le plus ancien ; est hautement virulent. Il semblerait donc que la virulence soit un caractère ancestral au sein du groupe B2 (lui-même étant le groupe le plus ancien au sein de l'espèce) et la présence de souches non-virulentes (sous-groupe VIII) serait un argument montrant que plusieurs événements de pertes (et/ou de ré-acquisition) de gènes de virulence ont eu lieu au cours du temps. Si le caractère virulent est présent dans le plus ancien des sous-groupes du plus ancien groupe de l'espèce, c'est que probablement ce caractère est soumis à une sélection positive. Pourtant, la virulence extra-intestinale est responsable de maladies graves (telles que les méningites néonatales et les septicémies) qui sont grevées d'une grande mortalité.

Pour quelle raison un caractère entraînant l'arrêt de la transmission d'une bactérie (en entraînant la mort de son hôte) pourrait-il être sélectionné ?

Pour résoudre ce paradoxe apparent, il faut envisager que les facteurs de virulence puissent conférer un avantage adaptatif dans des situations différentes de celles qui ont mis en évidence leur potentiel virulent. Donc que les facteurs dits de virulence soient aussi des facteurs de commensalisme.

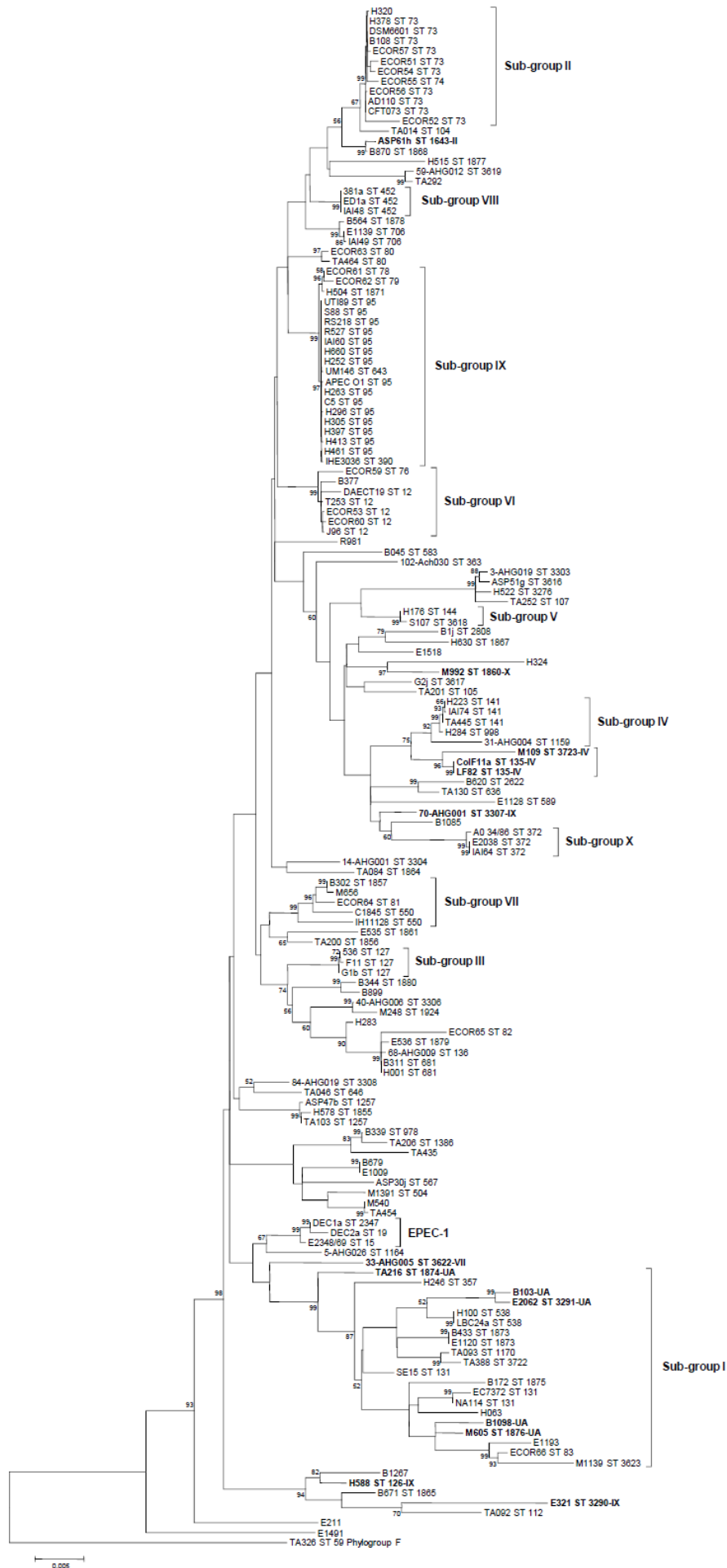


Figure 4 : Arbre phylogénétique mettant en évidence les 10 sous-groupes du groupe B2 de *E. coli* (d'après Clermont *et al.*, 2014).

Analyse de 153 isolats de *E. coli* appartenant au phylogroupe B2 en utilisant les séquences de 13 gènes (9807 nucléotides). L'arbre est enraciné à une souche de phylogroupe F (TA326). L'identité de chaque souche est suivie de sa désignation ST dans le schéma MLST de Achtman. UA correspond à des souches non-assignables.

III. Commensalisme et *E. coli*

Comprendre le chemin qui mène du commensalisme à la virulence nécessite de décrire au mieux ces 2 états. Si le second a suscité une importante masse de connaissances du fait de l'importance de *E. coli* en bactériologie médicale, le premier est, en revanche, moins bien décrit. C'est pourtant du fait du caractère ubiquitaire de l'espèce un territoire immense représentant sur terre environ 10^{20} individus (Whitman *et al.*, 1998) et concernant d'innombrables espèces d'hôtes ainsi que l'environnement inanimé (Tenaillon *et al.*, 2010, Savageau *et al.*, 1983, Solo-Gabriele *et al.*, 2000).

1). Distribution de *E. coli* dans la nature

E. coli est un des organismes modèles les mieux étudiés. Une somme considérable de connaissances en matière de génétique, de biologie moléculaire, de physiologie et de biochimie a été accumulée sur la souche K-12, une souche humaine commensale isolée en 1922 à Palo Alto. Toutefois, K-12 et les souches qui en dérivent sont loin de représenter la formidable diversité de l'espèce rencontrée dans le monde (Hobman *et al.*, 2007).

L'habitat naturel de l'espèce est l'intestin des vertébrés (Berg *et al.*, 1996, Gordon et Cawling, 2003). Bien que le microbiote intestinal contienne plus de 500 espèces bactériennes dont majoritairement des anaérobies, *E. coli* est le microorganisme aéro-anaérobie facultatif prédominant de la flore humaine (Berg *et al.*, 1996). *E. coli* est retrouvé dans le microbiote intestinal et à ce niveau, il constitue un compétiteur très performant. *E. coli* peut transiter dans l'eau et les sédiments, où ces bactéries sont excrétées à partir de leur habitat primaire, *E. coli* est d'ailleurs un indicateur de la contamination fécale de l'eau. Utilisant des calculs intuitifs, il a été estimé que plus de la moitié de la population de *E. coli* réside dans ces habitats secondaires (Savageau *et al.*, 1983). Mais au-delà de cette contamination, des études ont mis en évidence l'existence d'un saprophytisme véritable (Solo-Gabriele *et al.*, 2000).

L'adaptation de *E. coli* au tractus gastro-intestinal des animaux à sang chaud est considérée comme la plus importante source de sélection naturelle durant son évolution à partir d'un ancêtre commun avec *Salmonella*. *E. coli* a été initialement découvert et a été associé au microbiote intestinal des humains et des animaux. Ce qui est connu de *E. coli* (son abondance, sa distribution et les facteurs l'influençant) est basé sur des études d'isolats ou souches représentatives. Deux raisons font que les collections historiques d'isolats peuvent induire des biais. Tout d'abord, quand les isolats ont été rassemblés, les souches ont été caractérisées par leurs propriétés enzymatiques (ex. fermentation du lactose, production d'indole ou non utilisation du citrate) (Sneath 1973). Deuxièmement, les isolats n'ont pas été collectés de façon uniforme entre les différents habitats de *E. coli*. Une part disproportionnée de données a été rassemblée sur les souches de *E. coli* isolées chez les humains et les animaux domestiques. Il y a potentiellement un sous-échantillonnage de la diversité phénotypique et génotypique d'habitats non associés à l'hôte tels que l'environnement (Gordon *et al.*, 2002, Byappanahalli *et al.*, 2006, Walk *et al.*, 2007).

La caractérisation d'isolats de l'environnement change le point de vue sur l'histoire de *E. coli*. Ces observations sont en faveur de l'hypothèse que la population globale de *E. coli* n'est pas transmise directement entre hôtes par voie oro-fécale, mais plutôt par circulation des souches entre hôtes et environnement.

En situation commensale, les bactéries de l'espèce *E. coli* sont présentes aussi bien chez l'homme que chez de nombreux animaux. Dans le tractus gastro-intestinal, les souches de *E. coli* sont localisées de manière spécifique dans le côlon et le caecum des mammifères (Poulsen *et al.*, 1994). Elles résident dans le mucus qui tapisse l'épithélium intestinal et sont rejetées dans la lumière intestinale, avec ce même mucus dégradé, puis excrétées dans les selles (Tenaillon *et al.*, 2010). Le mucus définit une niche écologique à laquelle le métabolisme de *E. coli* s'est adapté (Freter *et al.*, 1983). C'est l'une des premières espèces bactériennes qui colonise le tube digestif après la naissance à partir des bactéries de la flore digestive de la mère (Bettelheim et Lennox-King, 1976). Elle atteint rapidement une très haute densité bactérienne (Mitsuoka et Hayakawa

1972; Penders 2006) avant l'expansion des bactéries anaérobies. L'amélioration des conditions d'hygiène à l'hôpital et dans les familles résidant dans les pays industrialisés a d'ailleurs réduit la colonisation précoce par *E. coli* (Nowrouzian *et al.*, 2003 ; Jauréguy *et al.*, 2004).

Les rapports entre *E. coli* et l'hôte sont un commensalisme où chacun des deux organismes bénéficie de l'interaction avec l'autre. *E. coli* tire ses nutriments de l'hôte (Conway 2004), ainsi qu'un environnement stable, une protection contre le stress et un moyen de transport et de dissémination. Toutefois, *E. coli* procure certains avantages à l'hôte en prévenant la colonisation par certains pathogènes comme les salmonelles, via la production de bactériocines entre autres ce qui se rapproche du mutualisme (Rastegar Lari *et al.*, 1990 ; Vollaard et Clasener 1994 ; Hudault *et al.*, 2001 ; Schamberger *et al.*, 2004).

2). Variabilité intra-hôte et inter-hôte:

De grandes variations dans la prévalence des groupes de *E. coli* sont retrouvées au sein des populations humaines.

Cependant, ces notions reposent sur une approche exclusivement qualitative puisqu'un seul clone par individu est étudié. Très peu est connu des aspects quantitatifs aussi bien en termes de quantité de corps bactériens viables (unités formant colonies, UFC) par gramme de selles que de répartition de ces souches dans les groupes phylogénétiques dans la flore de chacun des individus.

En étudiant jusqu'à 300 colonies par selles de certains individus, il a été estimé que chaque individu porte une souche dominante qui constitue plus de la moitié des colonies isolées, avec les autres souches présentes à des taux variables. Cette souche résidente est présente sur plusieurs mois ou années alors que d'autres souches peuvent transiter sur quelques jours ou semaines (Sears *et al.*, 1950, 1952, 1956, Bettelheim *et al.*, 1972, Caugant *et al.*, 1981). Les souches résidentes et en transit tendent à être identiques pour un individu donné. Plus récemment, il a été montré qu'il existe une diversité interhumaine de portage et que celle-ci est plus élevée en zone tropicale (Escobar-Paramo *et al.*, 2004, Skurnik *et al.*, 2008). Diverses études chez les membres d'une famille vivant ensemble ainsi que chez des partenaires sexuels ont montré que le portage de souches communes est dans ces cas statistiquement plus élevé que ce qui est observé communément entre adultes (Caugant *et al.*, 1984, Jonhson *et al.*, 2008). Dans les régions tropicales, les moins bonnes conditions d'hygiène alimentaire peuvent favoriser l'arrivée de nouvelles souches de *E. coli*.

Les expériences de colonisation intestinale chez les adultes par administration orale de souches de *E. coli* ont été décevantes (Cooke *et al.*, 1972) alors que la colonisation est aisée chez les nouveau-nés (Poisson *et al.*, 1986).

De plus, une souche de *E. coli* qui échoue à coloniser l'intestin d'une souris ayant un microbiote intestinal complet arrive à coloniser l'intestin d'une souris traitée par streptomycine

(Myhal *et al.*, 1982) ou d'une souris axénique (Freter *et al.*, 1983). De façon intéressante, cette souche tend à persister après la réimplantation d'un microbiote endogène (Freter *et al.*, 1983).

Ces résultats suggèrent une association forte entre la souche résidente et son hôte, association dans laquelle le reste du microbiote intestinal est impliqué.

Plusieurs études se sont intéressées à la diversité des souches commensales par différentes méthodes de discrimination et sont listées dans le tableau I.

Références	Nombre de sujets	Population concernée	Techniques	Nombre d'isolats par sujet
Sears <i>et al.</i> , 1950	4	4 adultes, Etats-Unis	O-typage	574, 1219, 102, 38
Caugant <i>et al.</i> , 1981	1	Un adulte suivi sur 1 an, Etats-Unis	MLEE	550
Lidin-Janson <i>et al.</i> , 1978	52	Enfants, Etats-Unis	Sérotypage et biotypage	10
Schlager <i>et al.</i> , 2002	13	Enfants, Etats-Unis	MLEE	28
Escobar-Paramo <i>et al.</i> , 2004	98	Adultes, France	PCR triplex	5 à 10
	93	Adultes, Guyane Française		
	46	Adultes, Bénin		
	28	Adultes, Colombie		
Nowrouzian <i>et al.</i> , 2005	70	Enfants, Suède	PCR triplex et RAPD	Différentes mophologies
Anderson <i>et al.</i> , 2006	5	Adultes, Etats-Unis	Ribotypage, résistances	15 à 20
Johnson <i>et al.</i> , 2007	630	Adultes, Etats-Unis	PCR triplex	1 à 3
Lautenbach <i>et al.</i> , 2008	49	Adultes, Etats-Unis	MLEE	25
Moreno <i>et al.</i> , 2008	67	Adultes, Espagne	PCR triplex	30
Moreno <i>et al.</i> , 2009	39	Adultes, Espagne	PCR triplex	30
Vollmerhausen <i>et al.</i> , 2011	59	Adultes, Australie	Caractères biochimiques et RAPD	28

Tableau I : Revue de la littérature des principales études portant sur l'estimation de la diversité des clones de *E. coli* dans les selles.

Ces études étaient lourdes, et typaient au mieux 30 colonies par selles, et de ce fait ne pouvait descendre en dessous du seuil de détection de 10% de la population de *E. coli*. Mais les premières études de Sears et de Caugant (Sears *et al.*, 1950, Caugant *et al.*, 1981) suggéraient une grande diversité de portage des souches commensales. Dans l'étude de Schlager (Schlager *et al.*, 2002), le changement de clone dominant sur plusieurs semaines suggérait soit sa disparition et l'acquisition de nouveaux clones, mais résulte plus probablement des variations dans le portage de celui-ci le faisant passer en dessous du seuil de détection et expliquant ainsi qu'il disparaisse puis réapparaisse au cours du temps. L'échantillonnage de quelques colonies prises au hasard ne semble refléter que la partie émergée de l'iceberg et une grande diversité de souches sous-dominantes restait encore à explorer.

3). Diversité des phylogroupes

Les facteurs impliqués dans la diversité phylogénétique des souches commensales dépendent essentiellement de l'environnement (Tableau II).

Son importance a été confirmée en comparant le microbiote intestinal de *E. coli* de sujets français de la métropole expatriés en Guyane française, de français résidant en métropole et de natifs de Guyane française. Le microbiote des expatriés était intermédiaire entre ceux des deux autres populations (Skurnik *et al.*, 2008). Le niveau d'hygiène et les habitudes alimentaires sont les principaux facteurs de distribution phylogénétique des groupes, plus que les conditions climatiques ou les caractéristiques génétiques de l'hôte ainsi que le suggère le changement radical de la proportion des groupes B2 et A en France durant les 20 dernières années (Gordon *et al.*, 2005). En région parisienne, il a été montré que le portage digestif de *E. coli* du groupe B2 est passé de 10% en 1981 à 37% en 2001 (Escobar-Paramo *et al.*, 2006). Cette tendance a été confirmée en Amérique du Nord, au Japon et en Australie, où, depuis les années 2000, les souches sont majoritairement de groupe phylogénétique B2 (Tableau III). Les groupes phylogénétiques retrouvés en zone tropicale sont différents de ceux en zone tempérée, avec une majorité de souches du groupe A (57%), suivi du groupe B1 (19%) (Escobar-Paramo *et al.*, 2004).

L'influence de la zone géographique a été également mise en évidence dans une étude Coréenne (province de Jeonnam, Corée du sud) où aucun isolat de phylogroupe B2 n'a été retrouvé ni chez les humains, ni chez les animaux domestiques alors que les *E. coli* B2 étaient présents dans les selles des animaux sauvages et dans les eaux des rivières (Unno *et al.*, 2009).

Tableau II : Répartition dans les groupes phylogénétiques des souches commensales d'origine humaine dans les années 1980 (Duriez *et al.*, 2001).

Population	Groupes				Nombre de sujets
	A	B1	B2	D	
France 1980	61	12,5	10,5	16	56
Croatie	35	32	19	14	57
Mali	24	58	2	16	55

Tableau III : Répartition dans les groupes phylogénétiques des souches commensales d'origine humaine dans les années 2000 (d'après Tenaillon *et al.*, 2010 et Lescat *et al.*, 2013).

Population	Groupes				Référence
	A	B1	B2	D	
Bénin	50	32,5	17,5	0	Escobar-Paramo 2004
Pakistan	47	18	12	23	Nowrouzian 2009
Guyane, France	64	20	3	13	Escobar-Paramo 2004
Brest, France	14,3	23,8	33,3	28,6	Escobar-Paramo 2004
Paris, France	29,6	11,1	37,1	23,3	Escobar-Paramo 2004
Bretagne (travailleurs du secteur tertiaire)	24	24	32	20	Escobar-Paramo 2004
Bretagne (Agriculteurs)	32	28	16	24	Escobar-Paramo 2004
France (2000)	25,5	21	29,5	24	Watt 2003
Colombie	57	3,5	25	14,5	Escobar-Paramo 2004
Bolivie	77	10	5	8	Pallecchi 2007
Suède	29	11	46	14	Nowrouzian 2005
Australie	19,5	12,5	45	23	Gordon 2005
Japon	28	0	44	28	Obata-Yasukoa 2002
USA	20,5	12,5	48	19	Zhang 2002
Guyane	54,3	23,5	3,7	13,6	Lescat 2013

De plus, au sein d'un même patient, le renouvellement des souches de *E. coli* est plus important en zone tropicale. Ceci a été démontré en comparant les souches commensales d'enfants pakistanais avec celles d'enfants suisses (Adlerberth *et al.*, 1998).

Des différences notables dans le portage d'un groupe phylogénétique donné sont retrouvées entre les humains et les animaux. Des caractéristiques telles que le régime, la morphologie intestinale et la masse corporelle semblent influencer sur la distribution des groupes phylogénétiques (Caugant *et al.*, 1984, Gordon *et al.*, 2003). Toutefois, la principale force façonnant la génétique de la population de *E. coli* chez l'animal est environnementale : c'est le degré de domestication (Escobar-Paramo *et al.*, 2006) (Tableau IV). Les animaux domestiques ont une proportion de souches B2 moindre que les animaux sauvages (30% chez les animaux sauvages *versus* 14% et 11% chez les animaux domestiques et de zoo), alors que la proportion de souches A augmente (de 14% chez les animaux sauvages à 27% et 26% chez les animaux domestiques et de zoo). Ces proportions ont été obtenues à partir de données de la littérature portant sur 1.154 animaux (Gordon et Cawling 2003, Escobar-Paramo *et al.*, 2006, Ochman et Selander, 1984, Baldy Chudzik *et al.*, 2008).

L'influence de la physiologie du tube digestif de l'hôte a par ailleurs, été retenue. Gordon et Cowling stipulent que les mammifères herbivores et omnivores présentent plus de souches des groupes B2 et D que les mammifères carnivores, les oiseaux et les reptiles et proposent l'hypothèse que ces souches ont les caractéristiques nécessaires pour coloniser des tubes digestifs plus complexes (Gordon et Cawling 2003). Ces observations étaient par ailleurs basées sur des résultats obtenus à partir de modèle *in vitro* de chémotats mimant ces différents tubes digestifs (Ballyk *et al.*, 2001). De même, bien que moins documenté, des niches différentes existent au sein de l'appareil digestif (Schierack *et al.*, 2008), les souches de phylogroupe A étant plus retrouvées dans le tractus digestif haut, alors que les souches B1 sont plus fréquemment isolées des selles (Dixit, 2004).

Tableau IV : Répartition dans les groupes phylogénétiques des souches commensales issues d'animaux domestiques et sauvages.

Caractéristiques des animaux	Groupes				Références
	A	B1	B2	D	
Animaux domestiques, France 1980	59	33	3	5	Escobar -Paramo 2006
Animaux domestiques, 2006	41	39	7	13	Escobar-Paramo 2006
Animaux sauvages, France 1980	68	9	4,5	18	Escobar-Paramo 2006
Animaux sauvages, 2006	32	24	20	24	Escobar-Paramo 2006
Animaux de zoo (USA) 1980	34	37,5	12,5	16	Ochman 1984
Animaux de zoo (USA) 2008	25	55	11	9	Baldy-Chudzik 2008
Animaux sauvages, mammifères, Australie	15,5	33	35	16,5	Gordon et Cawling 2003
Animaux sauvages, oiseaux, Australie	8	49	23	20	Gordon et Cawling 2003
Animaux sauvages, reptiles, Australie	12,5	75	7	5,5	Gordon et Cawling 2003
Animaux domestiques, Guyane	14,3	52,4	9,5	9,5	Lescat 2013
Animaux sauvages, Guyane	9	24,7	46,1	15,7	Lescat 2013

Ces études sur des souches commensales sur diverses populations humaines et animales ont démontré des associations entre l'espèce de l'hôte et la prévalence des principaux groupes phylogénétiques (Gordon et Cawling 2003, Escobar-Paramo *et al.*, 2006). Des facteurs d'alimentation, de climats, la période de recueil des selles, et l'origine géographique impactent aussi cette structure génétique comme l'ont démontré les larges études de Gordon et Cawling chez les animaux vertébrés en Australie (Gordon et Cawling 2003).

Afin d'éviter ces biais, une étude a été menée dans une zone géographique restreinte en Guyane Française et comparait les caractéristiques de souches humaines, et de souches issues d'animaux domestiques et d'animaux sauvages (Lescat *et al.*, 2013). Cette zone était d'accès difficile et les sujets bénéficiaient d'une prescription contrôlée d'antibiotiques. Les souches humaines appartenaient rarement au phylogroupe B2 (3.7%), et portaient peu de facteurs de virulence. A l'opposé les animaux sauvages portaient dans 46.1 % des cas des souches de phylogroupe B2, de STs rares, et riche en bactériocines, les souches d'animaux domestiques étaient intermédiaires.

4). Aspects quantitatifs du commensalisme

Les hôtes de *E. coli* ont des différences de masses corporelles, de régimes et de microbiotes. Ces différences influent sur la densité de l'espèce qui varie de 0 à 100 % et de plus de 6 ordres de magnitude parmi les espèces (Tenaillon *et al.*, 2010). Chez les humains, la prévalence de *E. coli* est de plus de 90% (Mitsukoa *et al.*, 1972, Penders *et al.*, 2006), mais elle est seulement de 56% chez les animaux sauvages, 23% chez les oiseaux et 10% chez les reptiles (Gordon et Cawling 2003). En Guyane Française, en étudiant dans une même unité géographique (village des Trois-Sauts et alentours) les selles d'Amérindiens, d'animaux domestiques et sauvages, la prévalence de portage baissait en fonction de la proximité à l'homme ainsi celle-ci était de 100% chez les humains, 64% chez les animaux domestiques, 45% chez les animaux sauvages (Lescat *et al.*, 2013). Par ailleurs, la prévalence de *E. coli* augmente avec la masse corporelle (Lescat *et al.*, 2013, Gordon et Cawling 2003).

La concentration par gramme de fèces varie chez les humains de 10^7 à 10^9 unités formant colonies (UFC) (Mitsukuo *et al.*, 1972 ; Penders *et al.*, 2006 ; Slanetz et Bartley 1957) et est plus basse chez les animaux domestiques où elle oscille entre 10^4 et 10^6 UFC. Il a été estimé que les mammifères et les oiseaux présentaient entre 10^6 et 10^9 UFC de *E. coli* par gramme de fèces dans le colon (Berg *et al.*, 1996, Gordon et Cawling 2003). Dans le travail comparatif entre microbiotes d'humains et d'animaux domestiques de Furet et coll. (Furet *et al.*, 2009), la concentration en *E. coli* par gramme de selles en Log UFC était de 7,7 ; 6,5 ; 6,4 ; 7,5 ; 7,8 et 9,2 pour les humains, les chevaux, les vaches, les chèvres, les moutons et les porcs respectivement.

E. coli appartient au microbiote intestinal qui consiste en plus de 500 espèces bactériennes qui totalisent 10^{10} à 10^{11} UFC par gramme de fèces. Les bactéries anaérobies sont prédominantes et dépassent *E. coli* d'un ratio de 100/1 à 10000/1 (Berg *et al.*, 1996). Comme nous l'avons vu, *E. coli* est parmi les premières espèces bactériennes à coloniser l'intestin chez les

nouveau-nés atteignant une densité de 10^9 par gramme de selles (Penders *et al.*, 2006) avant l'expansion des anaérobies (Syed *et al.*, 1970). Après l'âge de deux ans, cette densité se stabilise à environ 10^8 UFC par gramme jusqu'à ce qu'elle décroisse chez les plus âgés (Mitsukoa *et al.*, 1972).

Les travaux précédents abordent le commensalisme sous un angle uniquement qualitatif, en étudiant un ou plusieurs clones pris au hasard par sujet. Pourtant, l'obtention des données quantitatives précises permettrait une meilleure compréhension du commensalisme et de ses rapports avec la pathogénicité, et notamment l'effet de seuil dans les processus physiopathologiques de virulence mais aussi l'acquisition de résistances aux antibiotiques.

IV. Ilot de Haute Pathogénicité « HPI » et virulence

Dans la confrontation souvent ultime qui oppose le pathogène opportuniste à son hôte affaibli, la lutte pour la possession du fer est sans doute l'élément le plus déterminant.

Aux systèmes de captation du fer que possède l'hôte (ferritine, sidérophiline, protéines de l'inflammation), le pathogène oppose plusieurs systèmes de sidérophores. Aux moments critiques, le fond génétique de la souche de *E. coli* va, par le jeu de gènes régulateurs, conditionner l'échec ou la réussite de la confrontation.

1). Influence du fond génétique dans l'expression des gènes de virulence

Il a été proposé que les souches virulentes dérivent des souches commensales par différents types de mécanismes (Ochman *et al.*, 2000). L'acquisition de facteurs de virulence portés par des éléments génétiques mobiles (îlots de pathogénicité, plasmides tel par exemple le plasmide de virulence des *Shigella*) et la délétion de fragments d'ADN appelés "trous noirs" constituent deux mécanismes complémentaires pour l'évolution des pathogènes (Maurelli *et al.*, 1998).

Chez *E. coli* un lien clair a été établi entre la virulence extra intestinale et les groupes phylogénétiques B2 essentiellement et D, à un moindre degré. En étudiant 200 souches provenant d'infections extra intestinales et de selles de sujets sains, Goullet et Picard (1986) ont montré que près de 40 % des premières appartenaient au groupe phylogénétique B2, contre seulement 7 % des secondes. Ce pourcentage augmente à respectivement à 65%, 75%, et 84% des souches de septicémies (Johnson 1991; Jauréguy *et al.*, 2008), de méningites néonatales et d'urosepsis de nouveau-nés indemnes d'anomalies du tractus urinaire (Bidet *et al.*, 2007). Ces données épidémiologiques ont été confortées par l'utilisation d'un modèle de virulence extra-intestinale chez la souris qui testait la virulence intrinsèque des souches (Picard *et al.*, 1999). Quatre-vingt-deux souches de *E. coli* ont été testées avec ce modèle afin

d'en déterminer leur virulence extra-intestinale intrinsèque. La répartition de ces souches a ensuite été faite selon les groupes phylogénétiques principaux de l'espèce (A, B1, B2 et D). Cela a permis de montrer que le groupe phylogénétique B2 comportait un grand nombre de souches hautement virulentes (jusqu'à 100 % de mortalité sur le modèle murin de septicémie). Inversement, les souches des groupes A et B1 étaient souvent avirulentes.

D'autres travaux ont confirmé que le groupe B2 a le plus grand potentiel de virulence (Johnson *et al.*, 2006). Ainsi, de nombreuses études épidémiologiques et des résultats expérimentaux ont montré le rôle des souches appartenant au phylogroupe B2 dans la virulence ainsi que la présence chez ce phylogroupe de divers facteurs de virulence impliqués dans l'adhésion, la production de toxines, la capture du fer, et la protection de la cellule dans la virulence de la souche (Bingen *et al.*, 1998, Jonhson et Russo 2005, Picard *et al.*, 1999, Jonhson *et al.*, 2001, Bonacorsi *et al.*, 2003). Cependant, au sein de ce groupe, toutes les souches ne sont pas égales en termes de virulence. En effet, en inoculant des *E. coli* B2 d'origine pathogène extra-intestinale et commensales à des souris, environ 77% des souches sont très virulentes et dites « killer » (tuant plus de 8 souris sur 10), 16% sont dites « intermediate killer » (tuant entre 2 et 8 souris sur 10), mais qu'il existe aussi 7% des souches qui sont avirulentes et dites « non killer » (tuant moins de 2 souris sur 10), et ce, avec un pourcentage égal entre les deux populations de bactéries commensales et pathogènes étudiées. Comme expliqué auparavant, cette étude a permis de définir par MLST 9 sous-groupes au sein du groupe B2. Chez certains de ces sous-groupes toutes les souches sont virulentes chez la souris (I, III et IV), alors qu'au contraire, chez le sous-groupe VIII toutes les souches sont avirulentes (Le Gall *et al.*, 2007).

La distribution et la phylogénie de 17 gènes de facteurs de virulence (FV) ont été étudiées sur 98 souches de *E. coli* commensales et pathogènes intestinales et extraintestinales représentatives de l'espèce (Escobar-Paramo *et al.*, 2004). Les FV impliqués dans les

diarrhées de mécanisme aggrégatif ou entéro-pathogène étaient retrouvés et s'exprimaient dans tous les groupes phylogénétiques. Les FV impliqués dans les diarrhées hémorragiques ou toxiques étaient présents et s'exprimaient dans tous les groupes sauf les groupes B2 et D. Enfin, les FV associés aux ExPEC étaient identifiés dans tous les groupes phylogénétiques (sauf le groupe E). Cependant, chez les ExPEC, les proportions des groupes phylogénétiques B2 et D étaient plus importantes. En outre, lorsqu'on considérait les souches commensales, les souches du groupe B2 comportent davantage de facteurs de virulence ExPECs que les souches des autres groupes (Duriez *et al.*, 2001; Zhang *et al.*, 2002). Il apparaît donc que la présence et l'expression des FV expliquant les différentes pathologies engendrées par *E. coli* sont conditionnées par le fond génétique des souches dans lesquelles ils existent.

Comme l'ensemble de l'espèce *E. coli*, les différents pathovars ont une organisation clonale. Ils acquièrent cependant leurs FV par des transferts horizontaux. Cependant, ceux-ci n'affectent pas la structure clonale qui est conservée du fait que le fond génétique (qui est conditionné par la phylogénie de l'espèce) conditionne la présence et l'expression des FV. Ainsi, pour présenter une virulence intrinsèque, il ne suffirait pas d'être équipé de certains gènes de virulence, encore faudrait-il être du groupe phylogénétique B2 ou D.

Il apparaît donc clairement que la seule acquisition de facteurs de virulence n'est probablement pas suffisante pour permettre l'émergence de cette virulence. Au contraire, une interaction / adaptation étroite entre le génome receveur et les éléments génétiques acquis doit s'opérer pour permettre une intégration et une coopération efficaces entre ces éléments. L'ensemble de ces données met en relief le rôle essentiel de la régulation et du fond génétique pour l'expression de la virulence.

2). Îlots de pathogénicité

Les FVs peuvent être classés en grandes catégories fonctionnelles : toxines, systèmes de captation du fer, d'adhésion cellulaire (fimbriae, adhésines), de sécrétion, d'invasion cellulaire et protéines de régulation des activités des cellules de l'hôte (Jonhson et Russo 2005).

La majorité de ces VF sont rassemblés sur des éléments génétiques mobiles tels que des plasmides, ou sur des îlots génomiques appelés PAIs (Pathogenicity Associated Islands) (Hacker *et al.*, 1997) (ainsi les cinq systèmes de sécrétion des entérobactéries ont été retrouvés sur des PAI).

L'implication des PAIs dans la pathogénèse extra-intestinale a été démontrée de façon expérimentale (Lloyd *et al.*, 2009). En effet, l'étude des facteurs génétiques associés à la virulence, le plus souvent par le biais de mutants (Hacker *et al.*, 1983) a permis de montrer qu'ils étaient souvent physiquement liés sur des régions chromosomiques particulières. Il a été montré que ces groupements de gènes étaient retrouvés dans des clones différents d'*E. coli*, ce qui évoquait une acquisition par transfert horizontal (Low *et al.* 1984). Dans le même travail, il a été noté que l'insertion de ces groupements de VF se faisait grâce à de courtes séquences répétées faisant partie du groupement. Enfin, la démonstration que ces transferts horizontaux se faisaient le plus souvent à proximité d'un ARN de transfert, grâce à une machinerie de transposition génétique elle aussi présente (fonctionnelle ou non) sur le groupement de gènes a mené à la définition des îlots de pathogénicité (Hacker *et al.* 1990 ; Blum *et al.* 1994 ; Blum *et al.* 1995). La disposition en groupement de certains FV n'a ensuite cessé d'être confirmé (McDaniel *et al.* 1995 ; Mitsumori *et al.* 1998).

Les PAI sont caractérisés par les propriétés suivantes (Hacker *et al.* 1997; Dobrindt *et al.* 2004; Schmidt et Hensel 2004):

- pourcentage de G+C différent du reste du génome.

- biais de codon différent du reste du génome.
- ils comportent des groupements de gènes (« clusters ») impliqués dans la virulence de l'espèce.
- ils sont absents des souches non-virulentes de la même espèce ou d'espèces voisines.
- ils sont le plus souvent associés à une séquence codant pour un ARN de transfert (ARN_t).
- ils comportent des enzymes de mobilisation (transposases, intégrases) et des séquences d'insertion. Ces différents éléments peuvent être entiers et fonctionnels ou bien mutés ou tronqués et ne plus avoir d'activité.
- ils sont facilement mobilisables par transfert horizontal, dans une même souche, dans deux souches de la même espèce ou encore entre deux souches d'espèces différentes plus ou moins proches.

Des segments d'ADN ayant les mêmes caractéristiques ont été retrouvés sur des plasmides, des phages ou des transposons, certains PAI sont constitués de ces structures (ou une partie de ces structures) intégrées dans le chromosome.

Ils sont souvent encadrés par des courtes séquences répétées, probablement générées au moment de l'événement de recombinaison qui a permis leur arrivée dans un nouveau génome. Ils ont une structure souvent en mosaïque car ils sont composés de fragments d'ADN d'origines diverses, témoins de leur création en plusieurs étapes.

3). Le PAI IV ou HPI

La souche 536 a été isolée d'une infection urinaire (Hacker *et al.* 1983). L'étude de ses FV a mené à la découverte et à la définition des deux premiers PAI (Knapp *et al.* 1986 ; Marre *et al.* 1986). Deux PAI supplémentaires ont ensuite été définis par le biais de leur homologie avec des PAI connus d'autres souches pathogènes dont le HPI ou PAI IV.

En 1994, il a été démontré qu'une région particulière de l'ADN chromosomique de *Yersinia pestis* comportait un ensemble de gènes nécessaires à la virulence (Fetherston et Perry 1994). Peu après, il était montré que la délétion spontanée d'un large fragment de 102 kb d'ADN chromosomique était responsable de la perte de la virulence de la souche (Fetherston 1995 *et al.*, Carniel 1996). De plus, ce fragment était responsable de la synthèse d'un système de captation du fer, un sidérophore, la yersiniabactine décrite plusieurs années auparavant (Heesemann *et al.* 1993). Ce fragment chromosomique avait les caractéristiques d'un PAI et a donc été appelé HPI pour « High Pathogenicity Island » (Carniel *et al.* 1996) car il était essentiel à l'expression d'un phénotype hautement virulent. Il comporte un ensemble de 12 gènes impliqués dans la synthèse, dans la régulation du sidérophore et dans son transport via son récepteur (Rakin *et al.* 1999).

Toutefois, le HPI n'est pas limité aux *Yersiniae*. Ainsi, dans une étude de Bach et coll. (Bach *et al.* 2000), parmi 67 souches appartenant à 52 espèces d'entérobactéries, 13,4 % portaient le HPI. Ces isolats étaient des *E. coli*, *Citrobacter diversus* et des *Klebsiella* (*K. pneumoniae*, *K. rhinoscleromatis*, *K. ozonae*, *K. planticola* et *K. oxytoca*). La figure 5 représente les espèces et souches d'entérobactéries dans lesquelles le HPI a été identifié (Schubert *et al.* 1998). L'organisation du HPI chez *E. coli* ressemble à celle trouvée chez *Y. pestis* et *Y. pseudotuberculosis* et diffère de celle de *Y. enterocolitica* (Rakin et Heesemann 1995).

Cet îlot est fortement conservé entre les espèces avec des homologies comprises entre 98 et 100% entre les gènes du HPI de *Yersinia* et de *E. coli* (Dobrindt *et al.* 2002, Lesic et Carniel 2004, Schubert *et al.* 1998).

Chez *E. coli*, il est plus fréquent dans les souches responsables de septicémies (83 %), les souches entéro-aggrégatives (92 %) et les souches responsables d'infections urinaires (70 %). Il est également retrouvé chez 27% des *E. coli* producteurs de Shiga-toxine (STEC) (Karch *et*

al. 1999). Cependant, le HPI est aussi retrouvé dans 30% des isolats provenant de selles d'individus sains.

Dans l'étude de Clermont et coll (2001), la collection ECOR a été examinée, 93% des souches appartenant au phylogroupes B2 et D portaient le HPI contre 32% et 25% respectivement des souches des phylogroupes A et B1. Le HPI est retrouvé dans 100% des souches de méningites néonatales, dont certaines appartenait pourtant aux phylogroupes A et B1.

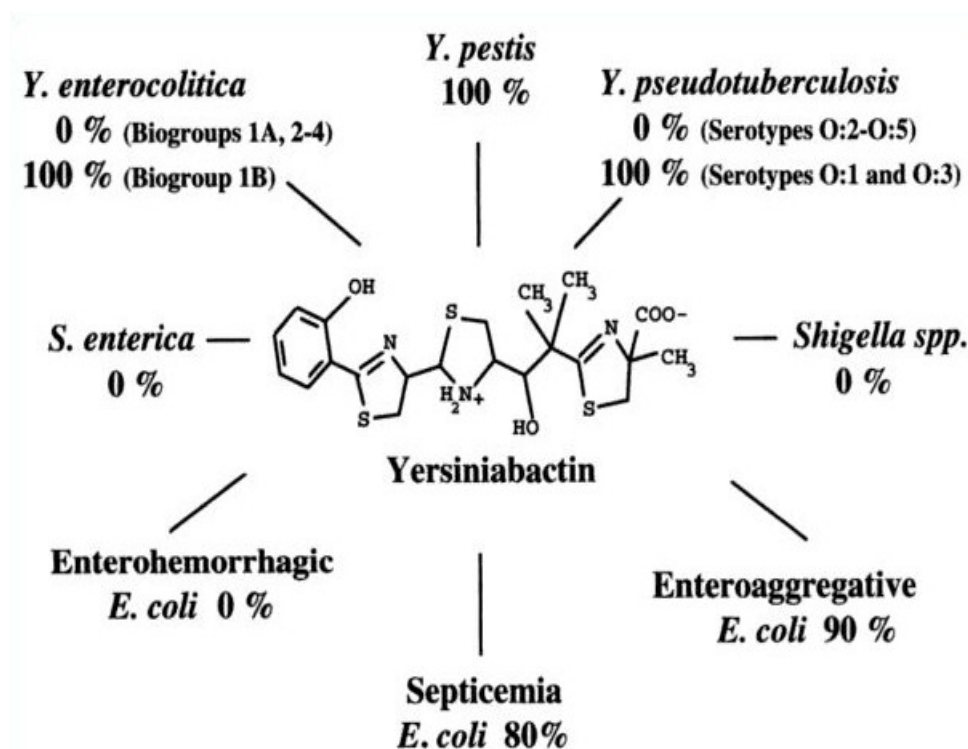


Figure 5 : Espèces et souches dans lesquelles le HPI (codant pour la yersiniabactine) de *Yersinia pestis* est présent d'après Schubert et al., 1998.

Le même îlot génomique peut donc être nécessaire à l'extrême virulence d'une espèce (*Yersinia pestis*), participer à la virulence de souches d'une autre espèce sans pour autant être nécessaire (*E. coli* entéro-agrégatif) ou encore être présent parmi les souches commensales d'un individu.

Le transfert horizontal du HPI au sein de l'espèce *E. coli* a été étudié (Schubert *et al.* 2009). Cinquante souches issues de la collection ECOR et représentatives de l'espèce ont été étudiées. Trente-sept portaient le HPI. La phylogénie (obtenue par MLST) de six gènes de ménage, six gènes du HPI ainsi que de fragments d'ADN en amont et en aval des HPI a été comparée. Cela a permis de montrer que le HPI n'était pas transmis que verticalement mais l'était également horizontalement à l'intérieur de l'espèce. Ce transfert portait sur de larges fragments d'ADN, comprenant le HPI ainsi que des régions chromosomiques flanquantes permettant la recombinaison. De plus, il a été montré que le transfert d'un HPI pouvait avoir lieu par conjugaison à l'aide du plasmide F. Le HPI peut donc être transmis par conjugaison puis disséminer au sein de l'espèce par recombinaison homologue.

Dans cette étude, le niveau de polymorphisme des gènes du HPI et des gènes de ménages ont été comparés ; les mutations synonymes considérées comme une horloge moléculaire sont rares dans le HPI comparées aux gènes de ménage plaçant pour une arrivée récente du HPI dans l'espèce *E. coli*.

La protéine TcpC est uniquement portée par les souches de *E. coli* de groupe phylogénétique B2, et présente des homologies avec les motifs du domaine Tir des Toll-like récepteurs. Elle est sécrétée par la bactérie, entre dans les cellules de l'hôte et interfère avec la signalisation de la voie pro-inflammatoire NFκB (Cirl *et al.* 2008). Le gène *TcpC* est situé à proximité du HPI sur le génome. La présence de TcpC est associée à la présence du HPI. TcpC présente une phylogénie incongruente avec la phylogénie des souches mais congruente avec celle du HPI. Ceci, avec la forte conservation du gène *tcpC*, plaide en faveur d'une introduction récente en bloc de ce gène avec le HPI par transfert horizontal dans l'un des points chauds majeurs de recombinaison dans le génome de *E. coli* (Schubert *et al.* 2010).

Le PAI IV est associé à l'ARN_t *asnT*. (Schubert *et al.* 1999, Karch *et al.* 1999) mais il peut être également inséré dans un autre site que l'*asnT* (Clermont *et al.* 2001). Ainsi, la souche ECOR31 porte un HPI unique de structure différente (ICE-type : integrative and conjugative element), et s'insère dans une copie de l'ARN_t *asnV* (Schubert *et al.* 2004).

Sous des conditions de croissance expérimentale, la majorité des PAIs peuvent être spontanément délétés du chromosome (Middendorf *et al.* 2004), toutefois le HPI chez la souche de *E. coli* 536 est stable ce qui est dû à la perte des séquences répétées flanquantes. Il manque en effet un DR (direct repeat) à l'extrémité 3' de l'îlot génomique (Dobrindt *et al.* 2002, Schubert *et al.* 1999). Ceci a déjà été rapporté pour la souche de *Y. enterocolitica* Ye8081 (Bach *et al.* 1999). Dans l'étude de Schubert *et al.* en 2009, il a été montré qu'à l'exception de la souche ECOR31, toutes les souches ECOR présentent une délétion unique de l'extrémité 3' du gène *fyuA* (codant pour le récepteur de la yersiniabactine) affectant le HPI et le génome de *E. coli* voisin « fixant » ainsi le HPI chez *E. coli*.

4). Structure du HPI :

Il a été montré que la transcription des gènes du HPI est fortement dépendante de la concentration en fer du milieu (Fetherston *et al.* 1996, 1999, Geoffroy *et al.* 2000).

Le HPI se compose d'une région conservée dite région core, et d'une région variable riche en A-T (Rakin *et al.* 1999). La région core comprend des gènes d'intégrases. Le gène d'intégrase (*int*) est souvent associé au génome phagique (intégrase analogue au gène d'intégrase du bactériophage P4) (Buchrieser *et al.* 1998).

La région core comprend également le gène du régulateur *ybtA* (Bearden *et al.* 1997, Buchrieser *et al.* 1999, Pelludat *et al.* 1998), le gène du récepteur de la yersiniabactine *fyuA* (Buchrieser *et al.* 1998, Carniel *et al.* 1996), les gènes de biosynthèse *irp1-5* (iron respressible protein) (*irp1*, *irp2*, *ybtU*, *ybtT* et *ybtE* chez *Y. pestis*) rassemblés dans un large opéron de 19-

kb (Gehring *et al.* 1998), et l'opéron polycistronique *irp6-9* (*ybtP*, *ybtQ*, *ybtX*, *ybtS* selon la nomenclature de *Y. pestis*) qui code pour les protéines Irp6 et 7 impliquées dans le transport Fe-yersiniabactine à travers la membrane cytoplasmique (YbtP et YbtQ chez *Y. pestis*), Irp9 (YbtS) qui est impliqué dans la synthèse du salicylate (Pelludat *et al.* 2003). Récemment, YbtX a été montré comme essentiel pour la captation de zinc chez *Yersinia pestis* (Bobrov *et al.*, 2014).

On compte ainsi quatre promoteurs correspondant à *ybtA*, *fyuA*, *irp1-5* et *irp6-9*. Les promoteurs de ces gènes contiennent une séquence palindromique conservée (Rakin *et al.* 1994), qui a été proposée comme site putatif de liaison d'YbtA (Fetherston *et al.* 1996) (Figure 6).

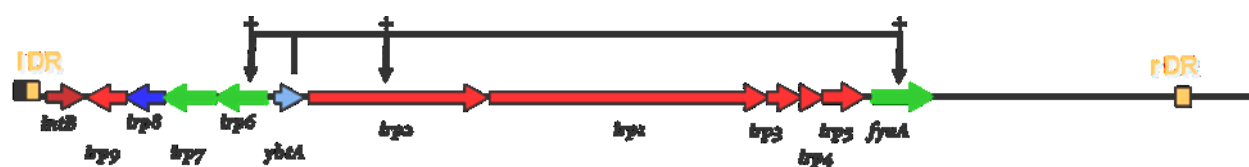


Figure 6 : Organisation génétique du HPI chez *E. coli*. La région gauche colorée correspond à la région core conservée, la région droite correspond à la partie variable riche en AT. Les gènes de fonction inconnue sont en bleu foncé, les gènes de transport en vert, les gènes de synthèse en rouge, les gènes de l'intégrase en marron, le gène de régulation en bleu clair avec les gènes régulés par lui indiqués par des flèches noires. Les extrémités correspondent à des directs repeats DR (Giuseppe Magistro, communication personnelle).

La présence d'un HPI chez *E. coli* ne signifie pas qu'il soit fonctionnel. Par exemple, chez les souches STEC, seules 60% exprimaient le récepteur FyuA. Ceci est dû, à des délétions partielles du gène *fyuA* (Schubert *et al.* 1998). La collection ECOR de 82 souches a

été analysée à la recherche des gènes du HPI et leurs fonctionnalités testée par un bioessai médié par un gène rapporteur (Schubert *et al.* 2000). De façon intéressante, le HPI était plus souvent fonctionnel dans les souches de phylogroupe B2 (92%) que dans celles des groupes A (71%) et B1 (45%), et montrait une forte association avec la létalité dans un modèle de virulence chez la souris (Schubert *et al.* 2002). Ceci pose la question du rôle du HPI en dehors de la production du sidérophore. Cet îlot génomique pourrait contribuer à augmenter la fitness de la souche, ou sa flexibilité métabolique ou son pouvoir pathogène en fonction du fond génétique sur lequel cet îlot arrive.

Une hypothèse évoquée est celle du régulateur *ybtA*. Chez *Yersinia*, il a été montré que la protéine yersiniabactine complexée à la protéine YbtA codée par le HPI est un régulateur positif de l'expression génique (Fetherston *et al.* 1996). Il est concevable que chez *E. coli* d'autres facteurs de virulence puissent être activés via le complexe YbtA-yersiniabactine.

YbtA, un régulateur transcriptionnel de type Ara-C, est le régulateur central de la production de yersiniabactine avec le régulateur Fur (Fe Uptake Regulation) (Anisimov *et al.* 2005). YbtA active la synthèse de yersiniabactine (opérons 1-5 avec le promoteur d'*irp2* en commun), le transport Fe-yersiniabactine (opéron 6-9 avec le promoteur d'*irp6*), la transcription de *fyuA*, et est répresseur de sa propre transcription. Comme les autres membres de la famille AraC, YbtA se fixe à l'ADN en complexe dimérique avec la yersiniabactine. Une étude récente a montré que la délétion de *ybtA* affectait le métabolome de *E. coli* et en particulier les voies de l'arginine (Lv *et al.*, 2011).

Les gènes du HPI ont une expression réprimée par le fer et par le régulateur ferrique principal Fur (Schubert *et al.* 2004, Fetherston *et al.* 1996). En effet, les quatre régions promotrices (*fyuA*, *irp2*, *ybtA* et *ybtP*) portent un site de fixation pour le répresseur Fur et ont leur expression réprimée en présence de fer (Carniel *et al.* 2001).

Les gènes *irp1* et *irp2* codent pour les protéines de synthèse de la yersiniabactine HMWP1(High-Molecular-Weight protein 1) et HMWP2 comme décrit précédemment pour *Yersinia* (Pelludat *et al.* 1998).

Le sidérophore yersiniabactine a été cristallisé (Miller *et al.* 2006) (Figure 7). Il est constitué d'un résidu salicylate, d'un résidu malonate, et des trois résidus de cystéine cyclique (Perry *et al.* 1999).

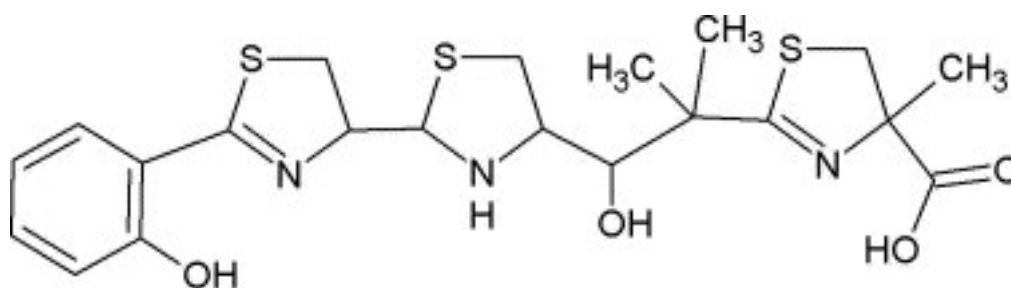
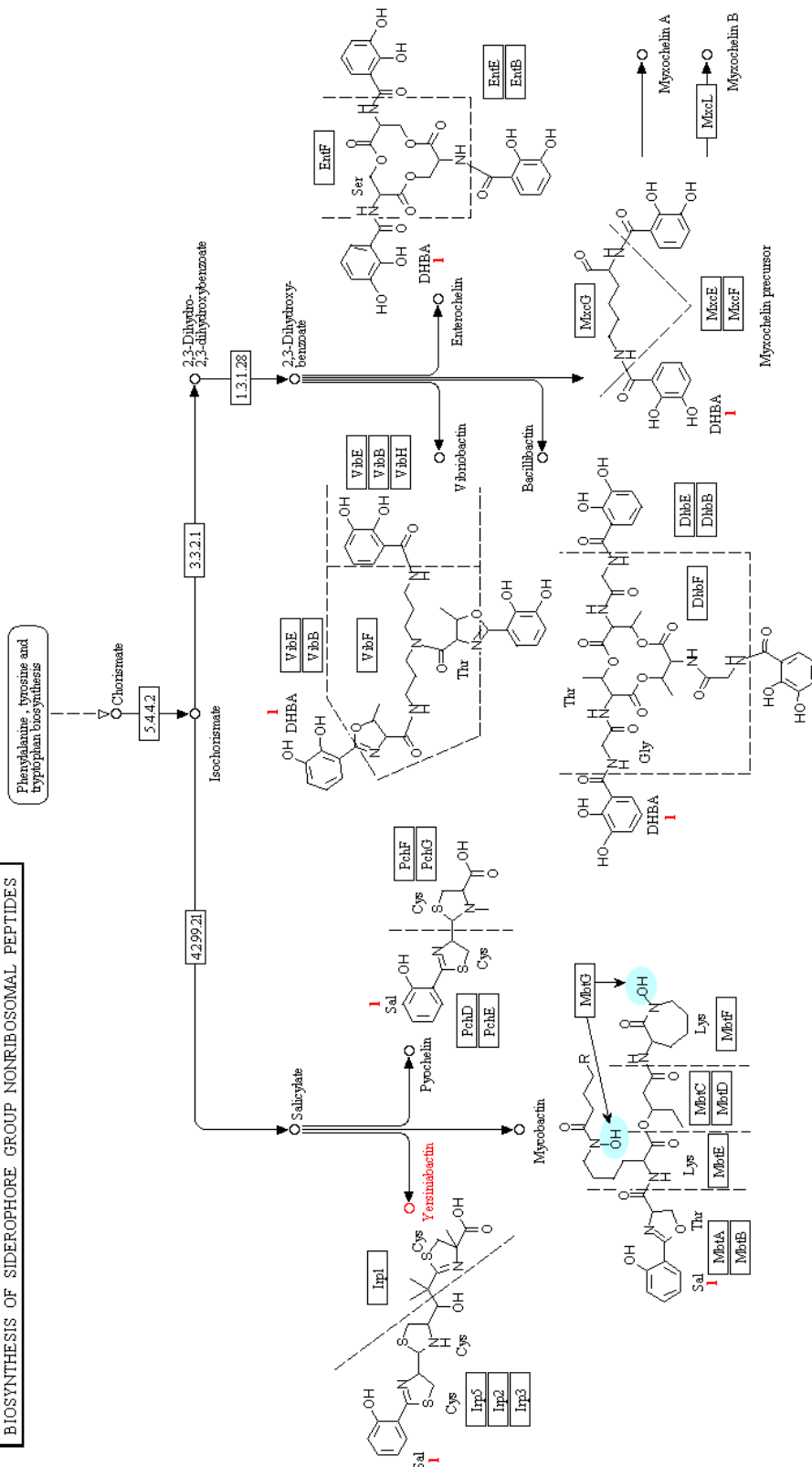


Figure 7 : Structure de la yersiniabactine d'après Miller *et al.*, 2006.

La synthèse du sidérophore se fait via un mécanisme non-ribosomique de synthèse polypeptidique (Figure 8). YbtS participe à la synthèse des dérivés salicylates. La synthèse de la yersiniabactine commence par YbtE qui adényle une molécule salicylate et la fixe sur HMPWP2 où deux molécules de cystéine sont alors ajoutées et rendues cycliques. Le substrat est alors transféré à HMWP1 où l'addition, la réduction et la méthylation d'un résidu malonate sont accomplies. Ces étapes sont suivies par l'addition et la cyclisation d'un dernier résidu cystéine. YbtU réduit la molécule (thiazoline en thizolidine) pour produire une molécule de yersiniabactine mature. Enfin, YbtT, une thioesterase putative, a une fonction d'« editing » ou de correction, ôtant les structures additionnelles ou aberrantes de la yersiniabactine (Bobrov *et al.* 2002).

BIOSYNTHESIS OF SIDEROPHORE GROUP NONRIBOSOMAL PEPTIDES



01053 1/24/12
(c) Kanehisa Laboratories

Figure 8 : Voies de biosynthèse des sidérophores du groupe des peptides non ribosomiques. Extrait de KEGG Kyoto Encyclopedia of Genes and Genomes (<http://www.kegg.jp/>).

On voit la proximité structurale entre la yersiniabactine et la pyochéline de *P. aeruginosa*.

Chez *Yersinia*, un gène *ybtD*, situé hors du HPI, a été identifié. Il n'était pas régulé par *ybtA* ou par le niveau de fer dans le milieu, mais était inducteur des gènes du HPI. Le gène *fuyA* code pour la protéine de 71.3 kDa extramembranaire FyuA (ferric yersiniabactin uptake), qui est un récepteur du complexe Fer-yersiniabactine et de la pesticine (Heesemann *et al.* 1993, Rakin *et al.* 1994, Pelludat 1998, Schubert *et al.* 2002). Appelée Psn chez *Yersinia*, elle fonctionne avec un transporteur ABC formé de deux perméases membranaires fixant l'ATP (YbtP et YbtQ chez *Yersinia*, correspondant à Irp6 et 7 chez *E. coli*) (Fetherston *et al.* 2010) (Figure 9). C'est un récepteur TonB-dépendant. La mutation d'un de ces trois gènes bloque la capture de fer via la yersiniabactine mais n'intervient pas dans la sécrétion de celle-ci.

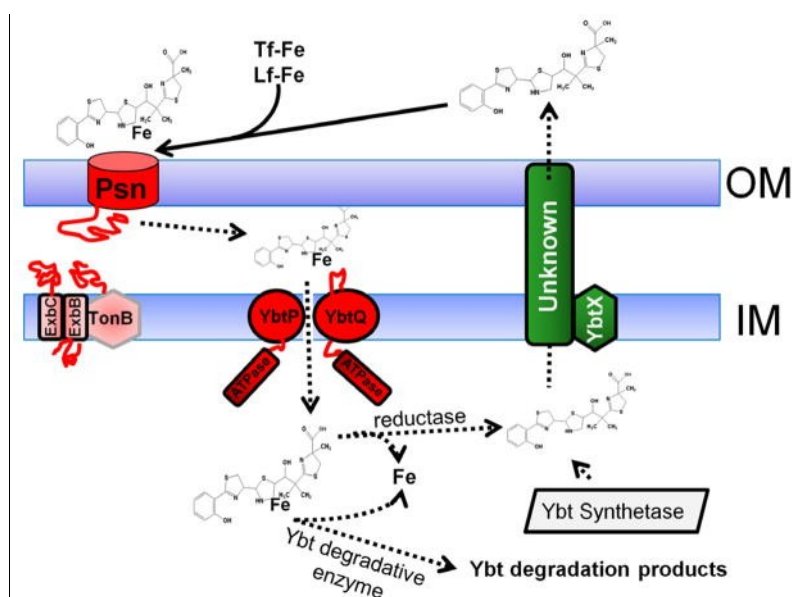


Figure 9 : Sécrétion et capture de la yersiniabactine chez *Yersinia*. La yersiniabactine (Ybt) est sécrétée par la cellule selon un mécanisme inconnu (boîtes vertes). YbtX est

impliqué dans la sécrétion mais n'est pas indispensable. Après sécrétion, la yersiniabactine peut capter le fer de la transferrine (Tf-Fe) et de la lactoferrine (Lf-Fe). Il semblerait que la yersiniabactine complexée au fer (Fe-Ybt) se fixe au récepteur Ton-B dépendant Psn (FyuA) puis transloque à travers la membrane externe. Une fois dans le périplasm, Fe-Ybt est transporté dans le cytoplasme par YbtP/YbtQ (Irp6/Irp7). Dans le cytoplasme, le fer pourrait être libéré du sidérophore par réduction de Fe^{3+} en Fe^{2+} ou par dégradation de Ybt. Les composants en rouge sont des éléments de transport, les lignes hachées indiquent des substrats ou des étapes non déterminées de façon expérimentale (Perry et Fetherston 2011).

5). Les autres sidérophores de *E. coli* :

Le fer est essentiel à la croissance bactérienne. Dans les hôtes mammifères, la quantité de fer libre est limitée et le fer se retrouve lié ou complexé à des protéines le séquestrant telles que la transferrine, la lactoferrine, l'hémoglobine et la ferritine; les bactéries étant ainsi dans un milieu pauvre en fer (Wandersman et Delepelaire 2004). En effet, la concentration en fer libre de 10^{-18} M est insuffisante pour permettre la croissance bactérienne (Carniel *et al.* 2001). Du fait de l'importance du fer dans une myriade de processus métaboliques, sa séquestration vis-à-vis des organismes infectieux est l'une des défenses innées contre l'infection (Miller *et al.* 2006).

La concentration de fer soluble est très basse dans l'urine et est connue pour limiter l'activité bactérienne (Roos et Klemm *et al.* 2006, Snyder *et al.* 2004). La captation du fer est critique pour la bactérie durant les phases de colonisation et d'invasion de l'hôte, d'où le développement par celle-ci de systèmes efficaces de captation du fer. Un système-type de captation à haute affinité du fer consiste en un chélateur de fer Fe^{3+} de bas poids moléculaire associé à son récepteur membranaire (Martinez *et al.* 1990). Un sidérophore est capable de chélater le fer de l'environnement ou de le prendre directement des protéines fixant le fer de l'hôte. Ces systèmes de captation du fer sont considérés comme des facteurs de virulence ou de fitness.

Chez *Yersinia*, le HPI est le seul système de sidérophores qui est critique pour sa virulence. D'autres systèmes d'acquisition du fer ont été identifiés depuis mais leur absence ne semble pas avoir d'impact sur la virulence (Gong *et al.* 2001, Kirillina *et al.* 2006).

Toutefois, *E. coli* possède d'autres sidérophores tels que l'entérobactine et beaucoup de souches produisent également l'aérobactine.

Au total, les ExPEC sont capables de produire jusqu'à 4 sidérophores : les catécholates entérobactine et salmochéline, l'hydroxamate aérobactine, et la yersiniabactine, sidérophore de type mixte (Henderson *et al.* 2009).

L'entérobactine (ou entérocheline) est un sidérophore de type catécholate qui est parmi les premiers sidérophores caractérisés et qui est produit par plus de 90% des isolats d'entérobactéries (Mokracka *et al.* 2004, Raymond *et al.* 2003, Visca *et al.* 1991). La souche K-12 ne produit qu'un seul sidérophore qui est l'entérobactine (O'Brien et Gibson 1970). L'entérobactine est un précurseur direct de la salmochéline, un autre sidérophore.

L'aérobactine est un sidérophore de type hydroxamate qui est produite par une plus petite fraction d'entérobactéries et qui a une affinité moindre pour le Fe³⁺ libre que l'entérobactine ou la yersiniabactine (Brock *et al.* 1991).

Alors que l'entérobactine est produite par presque toutes les souches de *E. coli*, les gènes codant pour l'aérobactine, la salmochéline et la yersiniabactine sont retrouvés plus souvent parmi les souches pathogènes (Kanamaru *et al.* 2003, Takahashi *et al.* 2006).

La souche probiotique de *E. coli* Nissle 1917 produit les quatre sidérophores (Valdebenito *et al.* 2006). Il est intéressant de noter que les souches responsables de bactériuries asymptomatiques, bien que dénuées des FVs des ExPEC ont un équipement complet en sidérophores, prouvant la nécessité de captation du fer pour réussir la colonisation urinaire (Roos et Klemm 2006).

In vitro, ces sidérophores ne sont pas produits dans les mêmes conditions de pH, de température et de source de carbone.

La production optimale de yersiniabactine et de salmochéline a lieu dans des conditions neutres à alcalines (pH 7.0 et 7.6, respectivement), alors que l'aérobactine est produite au maximum à pH acide (pH 5.6). Des mutants délétés de la yersiniabactine mis en compétition de culture avec la souche sauvage étaient supplantés par celle-ci à pH neutre et alcalin. Il

semblerait que la yersiniabactine soit par contre instable dans des conditions acides (Haag 1993 *et al.*).

Cet arsenal de sidérophores permet à la souche de s'adapter à des pH variables. Ainsi le pH de l'urine humaine varie entre 4.6 et 8.0 (Valdebenito *et al.* 2006).

La constante d'affinité (pK) avec le Fe^{3+} de la yersiniabactine est de 36,6, et est la plus élevée comparée à celles de Fe^{3+} -entérobactine (35,5) et de Fe^{3+} -aérobactine (23,3) (Perry *et al.* 1999).

Le sidérophore hydrophobe entérobactine se fixe à l'albumine humaine (Konopka et Neilands, 1984). De plus, il a été démontré que l'entérobactine était séquestrée par la lipochaline 2 dans un modèle d'infection péritonéale chez la souris (Flo *et al.* 2004). La lipochaline est stimulée par la reconnaissance des bactéries par les récepteurs Toll-like, elle est surexprimée au niveau des muqueuses en cas de colonisation bactérienne, d'où l'intérêt pour la bactérie à posséder d'autres sidérophores que celle-ci (Nelson *et al.* 2005, Lawlor *et al.* 2007). La salmochéline, forme glycosylée de l'entérobactine résiste à l'inactivation par la lipochaline 2, et représente une stratégie d'échappement immunitaire (Fischbach *et al.* 2006).

Ces sidérophores ont été étudiés pour leur importance durant l'infection.

La virulence des souches de méningites néonatales C5 et S88 est atténuée quand il manque IroN, récepteur de la salmochéline (Nègre *et al.* 2004, Peigne *et al.* 2009) dans un modèle de méningite du rat nouveau-né. Il semblerait que la salmochéline soit indispensable dans la phase bactériémique de la méningite à *E. coli*. L'aérobactine ne semblait pas critique dans ce modèle.

Par contre, les souches de *E. coli* produisant l'aérobactine sont plus virulentes dans un modèle d'infection sous-cutanée chez la souris (Demir et Kaleli 2004).

Les gènes d'acquisition du fer sont les plus surexprimés chez la souche de *E. coli* CFT073 lors d'une infection expérimentale du tractus urinaire chez la souris (Snyder *et al.* 2004). De

même, ces systèmes sont fortement exprimés dans des souches cliniques d'infection urinaire prélevées directement chez des femmes atteintes de cystites (Hagan *et al.* 2010), confirmant l'importance des sidérophores dans la pathogénicité.

Les mutants d' IroN récepteur de la salmochéline (souche CP9), d' IutA récepteur de l'aérobactine (souche CFT073), ou d' Iha récepteur de l'entérobactine (souche CFT073) ne semblaient pas diminués dans leur capacités de colonisation du tractus urinaire mais, dans tous les cas, s'ils étaient mis en compétition avec la souche sauvage, celle-ci les supplantait rapidement (Jonhson *et al.* 2005, Russo 2002, Torres *et al.* 2001). Les mutants du récepteur de la salmochéline IroN étaient affectés dans leur capacité d'invasion des cellules urothéliales (Feldmann *et al.* 2007) et dans la colonisation rénale (Russo *et al.* 2002). Les souches de *E. coli* déficientes en aérobactine et en entérobactine (doubles mutants de CFT073) montraient une moindre dissémination dans le rein dans un modèle d'infection urinaire chez la souris, toutefois les mutants déficients en aérobactine seule ou entérobactine seule ne montraient pas de défauts de colonisation ou de dissémination (Torres *et al.* 2001), d'où la notion de redondance des sidérophores (Garcia *et al.* 2011).

Tous les récepteurs des sidérophores sont fortement exprimés à la surface bactérienne pendant l'infection (Braun *et al.* 2005, Hagan et Mobley 2007). Les gènes des récepteurs sont les plus surexprimés durant l'infection urinaire (expression *in vivo* élevée). Ces récepteurs sont induits pendant la croissance dans l'urine humaine, exposés à la surface bactérienne, antigéniques et ayant une spécificité de pathogène, ce qui en fait des candidats vaccinaux idéaux. Ces récepteurs (FyuA, IroN et IutA) ont d'ailleurs été utilisés comme cibles vaccinales pour prévenir les infections dues aux ExPEC (Wieser *et al.* 2010, Durant *et al.* 2007), et notamment les infections urinaires (Alteri *et al.* 2009).

6). Virulence expérimentale liée au HPI :

Les souches hautement pathogènes de *Y. pestis*, *Y. pseudotuberculosis* et *Y. enterocolitica* biotype IB possèdent le HPI. Cet îlot est nécessaire à la virulence de *Yersinia* (Carniel *et al.* 1999, Bearden *et al.* 1997, Brem *et al.* 2001). Il n'a jamais été détecté dans les souches de moindre virulence ou avirulentes de *Yersinia* (De Almeida *et al.* 1993, Carniel *et al.* 2001). On rapporte dans la littérature des cas d'infections systémiques sévères à *Yersinia enterocolitica* peu pathogènes chez des hôtes présentant des excès en fer (hémochromatose, thalassémie...) prouvant l'importance de la capture du fer pour la dissémination de la bactérie.

Chez *Yersinia*, le système yersiniabactine est capable de déplacer le fer de la transferrine et de la lactoferrine. Son intégrité est indispensable à la virulence de *Y. pestis* dans un modèle d'infection par voie sous-cutanée chez la souris. Ainsi cette virulence est atténuée que ce soit avec des mutants de *psn* (*fyuA*), de *ybtP* (*irp6*) ou d'*irp2*, gène de synthèse (Fetherston *et al.* 2010). Il est également capital dans la survenue des formes pulmonaires de la peste. De façon intéressante, dans le modèle pulmonaire, les souches produisant la yersiniabactine mais incapable de l'utiliser en tant que sidérophore (*psn* mutants) étaient plus virulentes que les mutants des gènes de synthèse (Fetherston *et al.* 2010). Deux hypothèses étaient proposées par les auteurs : soit le sidérophore a un rôle de signalisation et de petites quantités de celui-ci suffisent en entrant dans la bactérie via des voies alternatives, tel que le sidérophore pyoverdine chez *Pseudomonas aeruginosa* qui régule sa propre production ainsi que celles de VF₃ (Lamont *et al.* 2002), ou bien le sidérophore affecte l'immunité innée, comme l'entérobactine qui est cytotoxique pour les cellules T (Hileti *et al.* 1995). La pyochéline qui ressemble structurellement à la yersiniabactine, peut produire des radicaux hydroxyles et causer des lésions de l'endothélium et de l'épithélium pulmonaire (Britigan *et al.* 1997, Britigan *et al.* 1994, Coffman *et al.* 1990).

Il est également à noter que pour *Y. pestis*, l'atténuation de virulence via le HPI est spécifique de la voie sous-cutanée, les mutants administrés par voie intraveineuse étaient

complètement virulents (Bearden *et al.* 1997). Les auteurs concluaient que la yersiniabactine était critique dans la phase lymphatique précoce de la peste bubonique mais pas dans la phase septicémique (Perry *et al.* 2011). En utilisant la *gfp* comme gène rapporteur dans un modèle de souris, il a été démontré que la production de yersiniabactine est élevée chez les souris infectées au niveau de la rate et des plaques de Peyer, mais basse dans la lumière intestinale (Jacobi *et al.* 2001).

Le transfert des gènes du HPI à une souche de *Y. enterocolitica* O:9 de faible pathogénicité, suffit à augmenter sa virulence chez la souris (Pelludat *et al.* 2002).

De la yersiniabactine purifiée a baissé la production de ROS (reactive oxygen species) par des monocytes humains et des polynucléaires neutrophiles *in vitro* (Paauw *et al.* 2009).

Récemment, il a été montré que la yersiniabactine fixait les ions cuivre, et permettait ainsi à *E. coli* de survivre dans les phagocytes en les protégeant de la bouffée oxydative (Chaturvedi *et al.* 2012, Chaturvedi *et al.* 2014).

Chez *Klebsiella pneumoniae*, la yersiniabactine joue un rôle critique dans l'acquisition du fer durant l'infection pulmonaire dans un modèle d'inoculation intranasale chez la souris (Lawlor *et al.* 2007). Cette souche produisait l'entérobactine et la yersiniabactine mais pas l'aérobactine. La yersiniabactine dans cette étude n'était pas nécessaire à la croissance *in vitro* en milieu carencé en fer, n'affectait pas la phase précoce d'infection pulmonaire chez la souris (les 24 premières heures), ce qui plaide en faveur d'absence d'impact sur la fitness ou de modulation de l'immunité innée, par contre à 96h, il y avait un défaut majeur de croissance de la bactérie déficiente en yersiniabactine *in vivo* dans le poumon.

De plus, il a été montré que dans les voies aériennes, il y a une production majeure de lactoferrine qui pourrait servir de donneur de fer pour catalyser une bouffée oxydative (Ward et Conneely 2004). Le sidérophore yersiniabactine pourrait donc avoir un rôle capital à ce

niveau. En effet, la yersiniabactine n'est pas séquestrée par la lipochaline 2, élément clé de défense des voies respiratoires contre *K. pneumoniae* (Bachman *et al.*, 2011).

Des études se sont intéressées au rôle du HPI dans les souches de *K. pneumoniae* responsables d'abcès hépatiques communautaires (souches porteuse du gène *magA*, hypermuqueuses, résistantes au sérum, résistantes à la phagocytose intra-macrophagique et hypervirulentes chez la souris (Fang *et al.* 2004). Ces souches de *K. pneumoniae* sont plus souvent porteuses de HPI que celles n'envahissant pas les tissus (non responsables d'abcès pyogènes hépatiques).

Des études ont montré la présence d'un HPI mobilisable porté par un ICE (élément intégratif et conjugatif) , avec conjugaison possible à d'autres souches, expliquant sa dissémination.

Il a été montré que le HPI est critique pour la phase bactériémique dans un modèle de septicémie chez la souris chez les souches de *E. coli* IA151 et IA152 (ces souches ne produisaient pas certains FV comme l'hémolysine et l'aérobactine) (Schubert *et al.* 2002), mais que le HPI n'avait pas d'impact dans un modèle de méningite néonatale chez le rat nouveau-né (souche C5) (Nègre *et al.* 2004), ni dans un modèle d'infection urinaire ascendante chez la souris (souche 536) (Brzuszkiewicz *et al.* 2006), ni dans un modèle de septicémie chez la souris (souche 536) (Tourret *et al.* 2010).

Tous les gènes du HPI étaient surexprimés dans des modèles de biofilms mais pas ceux des sidérophores, aérobactine et entérobactine (Hancock et Klemm 2007). Le gène *fyuA*, qui code pour le récepteur de la yersiniabactine, est l'un des gènes les plus exprimés dans le biofilm dans un modèle de chambres à flux, mimant des cathéters (63 fois dans la souche *E. coli* uropathogène VR50). La protéine FyuA est critique pour la formation de biofilm dans des milieux pauvres en fer comme l'urine humaine. Les mutants *fyuA* montraient une formation aberrante de biofilms et les bactéries devenaient filamenteuses (Hancock *et al.* 2008). L'introduction d'un plasmide codant pour FyuA dans les souches 536 et NU14 a augmenté

leur capacité à former du biofilm. Ce gène a été montré surexprimé 25 fois dans le tractus urinaire humain (Roos and Klemm, 2006).

7). Modèle *Dictyostelium discoideum*

L'amibe *D. discoideum* a été utilisée comme modèle de virulence expérimentale dans l'étude de la virulence conférée par le HPI.

Les amibes sont des protozoaires prédateurs qui chassent les bactéries. Elles exercent de ce fait un contrôle important sur les populations bactériennes et leurs densités. Elles sont fréquemment retrouvées dans les écosystèmes résultant de l'activité humaine tels que les systèmes de climatisation, l'eau du robinet, et les usines de traitement des eaux usées où elles se nourrissent de biofilms microbiens. De nombreuses espèces bactériennes ont développés des mécanismes de résistance à la digestion par les amibes. Ainsi *Legionella pneumophila* (Solomon *et al.*, 2000) est capable de mécanisme d'échappement à la digestion par la phagocytose et c'est cette capacité de multiplication au sein de l'amibe et à infecter les macrophages qui est à l'origine de sa virulence chez l'homme (légionellose). Ainsi, sa capacité d'adaptation à l'amibe est un facteur de virulence chez l'humain.

Cependant, très peu d'études ont porté sur les relations hôtes pathogènes amibes-bactérie *E. coli* (Adiba *et al.*, 2010). *D. discoideum* est une cellule phagocytaire, qui vit dans les sols et se nourrit de microorganismes comme les bactéries et les levures (Bonner *et al.* 1950, Steinert *et al.*, 2003). Les bactéries une fois digérées, lui apportent les nutriments nécessaires à sa survie et sa multiplication (Figure 10). Cette capacité à se nourrir des bactéries a été utilisée depuis longtemps en laboratoire pour cultiver les amibes (Raper and Smith, 1939). *D. discoideum* représente donc un prédateur naturel pour les bactéries environnementales (dont *E. coli*).

Les souches de *D. discoideum* initialement isolées dans l'environnement ont été clonées pour obtenir des souches de laboratoire dites axéniques, c'est-à-dire capables de croître en l'absence de bactéries (Urushihara *et al.*, 2009). Ces souches sont capables d'ingérer le milieu de culture par macropinocytose. Dans un milieu nutritif, les cellules de *D. discoideum* se

divisent en moyenne 2 fois par jour (12 h). Les cellules peuvent être cultivées sous forme adhérente dans une boîte de Pétri ou en culture liquide. Si l'on dépose une amibe *Dictyostelium* sur une couche uniforme de bactéries sur milieu agar, et après consommation du milieu nutritif, cette amibe ingère les bactéries et se multiplie, ce qui se manifeste après quelques jours d'incubation par l'apparition d'une zone circulaire dépourvue de bactéries, appelées plage de phagocytose, comparable à une plage de lyse de bactériophages. Lors d'une carence en nutriments, *D. discoideum* entame un cycle de développement multicellulaire complexe (il ne s'agit pas d'un syncytium), cette agrégation est contrôlée par l'AMPc et produit des fructifications ou spores (Otto *et al.*, 2001) (Figure 11).

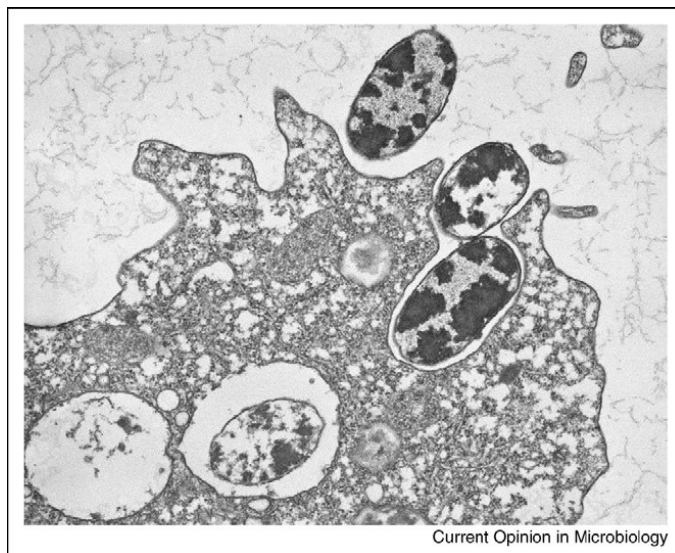


Figure 10 : Photographie en microscopie électronique du processus d'internalisation par l'amibe. Internalisation de la bactérie *Klebsiella* montrant les différentes étapes de la phagocytose par l'amibe: adhésion de l'amibe à la bactérie, formation d'un phagosome englobant la bactérie et digestion des bactéries internalisées (Cosson et Soldati 2008).

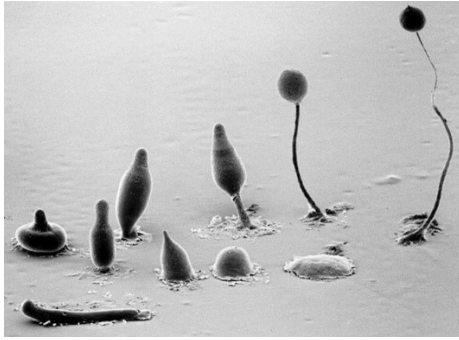


Figure 11 : Photographie en microscopie électronique des différents stades de développement de *Dictyostelium discoideum* d'après M.J. Grimson et R.L. Blanton (www.dictybase.org).

Il y a une compétition entre l'amibe et les bactéries, les facteurs de virulence bactériens leur permettant de se défendre contre l'amibe (Hilbi *et al.*, 2007). Les souches virulentes sont capables de résister à la phagocytose par *D. discoideum*, définissant ainsi le phénotype « Grazing R » (résistant au broutage par l'amibe) par opposition au phénotype « Grazing S » (sensible au broutage par l'amibe). *D. discoideum* est un bon modèle hôte pour étudier la virulence des bactéries. Ainsi, *D. discoideum* a permis l'étude et l'identification de nouveaux facteurs de virulence de *P. aeruginosa* (Alibaud *et al.*, 2008, Lima *et al.*, 2010).

Ce protozoaire du sol possède de nombreux gènes homologues à ceux des eucaryotes supérieurs et partage de nombreuses fonctions physiologiques avec les cellules de mammifères, ce qui en fait dans notre cas un modèle pour l'étude de la phagocytose (Ruuper et Cardelli, 2001, Chen *et al.*, 2007). Son génome est entièrement séquencé depuis 1995. Il mesure 34 Mb, contient six chromosomes codant approximativement pour 12 500 protéines. Une plate-forme d'échanges et de bases de données via le site DictyBase qui rassemble la communauté scientifique travaillant sur le *Dictyostelium* (www.dictybase.org).

D. discoideum émerge depuis une quinzaine d'années comme un véritable modèle d'étude des relations hôte-pathogène. De nombreux pathogènes ont été étudiés chez *D. discoideum* depuis : *Legionella pneumophila* (Hagele *et al.* 2000, Solomon *et al.* 2000), *Aeromonas* (Froquet *et al.* 2007), *M. avium* et *M. marinum* (Skriwan *et al.* 2002, Solomon *et al.* 2003), *P. aeruginosa* (Cosson *et al.* 2002), *Vibrio cholerae* (Pukatzki *et al.* 2006) et *K. pneumoniae* (Benghezal *et al.* 2006).

Il est à noter que la virulence de *Pseudomonas* dépend largement des toxines sécrétées qui tuent les amibes, pendant que la virulence de *Klebsiella* repose sur sa capacité à résister à la digestion intra-macrophagique.

Comparé à d'autres modèles de virulence, *D. discoideum* est certainement le plus simple d'utilisation grâce à son temps de génération extrêmement court (12 h environ), sa très petite taille, et sa culture simple et peu onéreuse. *D. discoideum* présente une bonne reproductibilité des résultats. Ceci est dû au fait que l'organisme est unicellulaire et assimilable à un macrophage et que les tests sont effectués sur une population cellulaire clonale. Chez la souris, la variabilité des résultats est une conséquence de la singularité de chaque individu d'une part, mais aussi de la complexité de cet organisme avec activation d'un système immunitaire beaucoup plus évolué. Autre avantage, *D. discoideum* est un des seuls modèles de virulence à offrir la capacité de tester la virulence bactérienne à grande échelle et dans un temps réduit (Alibaud *et al.*, 2008), ce qui permet d'étudier rapidement un grand nombre de souches.

Par contre, ses limites sont l'absence du système NFkB et des caspases, et la croissance à 27°C, ce qui empêche d'étudier des phénomènes physiologiques se produisant uniquement à 37°C.

INTRODUCTION DE LA PARTIE

EXPERIMENTALE

Nous avons vu dans la partie bibliographique que les progrès des techniques de biologie moléculaire ont permis de préciser la structure génétique de l'espèce *E. coli* et de définir assez clairement l'organisation des clones en groupes et sous-groupes phylogénétiques au sein de cette espèce.

Cependant, si une souche donnée peut être précisément assignée à un phylum déterminé, ce qui dans certaines circonstances est d'une grande contribution en termes épidémiologiques comme pour l'émergence du clone de *E. coli* ST131 responsable d'une véritable pandémie; en revanche, le positionnement des groupes et sous-groupes phylogénétiques dans les différentes niches écologiques de l'espèce et dans le répertoire des situations physiopathologiques des infections à *E. coli* est loin d'être encore défini.

Dans ce but, les stratégies antérieures consistaient essentiellement à typer un clone bactérien au hasard sur un milieu d'isolement et à l'assigner à un phylum puis de comparer la fréquence de ce phylum dans différentes situations épidémiologiques (pathogénicité, commensalisme, saprophytisme, espèce d'hôte etc). Ces stratégies basées sur l'hypothèse d'un « clone dominant » n'apportent pas d'information sur l'existence de clones minoritaires chez les souches commensales et les souches virulentes et ne permettent pas de proposer des profils d'abondance de populations bactériennes. Or, ces clones minoritaires existent tant en situation de commensalisme (Sears *et al.*, 1950, Bettelheim *et al.*, 1972, Lautenbach *et al.*, 2008, Moreno *et al.*, 2009, Schierack 2009, Smati *et al.*, 2013) que de pathogénicité (Levert *et al.*, 2010). Les connaître, en déterminer leurs proportions dans des panels statistiquement significatifs de selles d'individus strictement sélectionnés pour être en situation de commensalisme et de selles de différentes espèces animales choisies pour la variabilité de leur taille, de leur régimes alimentaire et de leur mode de vie est l'objet de la première partie de ce travail expérimental.

Etudier l'adaptation c'est rechercher la capacité des clones de *E. coli* à survivre et à se développer dans des milieux aussi différents que l'intestin, l'arbre urinaire, la bile, le sang, le liquide céphalo-rachidien de l'homme et de divers espèces animales.

Pour cela, ces clones doivent présenter des propriétés de mobilité, d'échappement au système immunitaire (complément, phagocytose) mais également des propriétés métaboliques dans des situations de stress ou de concurrence extrême.

Parmi les différents métabolites nécessaires au développement bactérien, le fer est un facteur de croissance crucial. Pour s'en procurer les clones *de E. coli* vont devoir concurrencer les tissus où ils se développent et leurs équipements en système de captation du fer. Cela explique la dotation des clones de *E. coli* en divers sidérophores.

L'un d'entre eux, la yersiniabactine, est largement répandu dans les différents phylums de l'espèce et ses effets semblent être cruciaux dans la physiopathologie des infections naturelles (Clermont *et al.*, 2001) et dans différents modèles de pathologie expérimentale (Schubert *et al.*, 2002).

La seconde partie du travail expérimental a consisté à étudier l'influence de l'environnement génétique de différents phylums de l'espèce en comparant les effets de plusieurs composantes de la yersiniabactine : le sidérophore, son transport, sa régulation à l'aide d'un panel de souches appartenant à plusieurs sous-groupes du groupe phylogénétique B2, dans 2 modèles de pathologie expérimentale, d'une part la septicémie chez la souris et d'autre part les rapports de *E. coli* avec l'amibe sociale *D. discoideum*.

ARTICLE N°1

I. Introduction de l'article n°1

Quel que soit son origine, *E. coli*, comme les autres entérobactéries, s'est adapté au cours de l'évolution au tube digestif des animaux à sang chaud qui constitue son principal habitat.

Chez l'homme, *E. coli* est l'entérobactérie la plus fréquemment mise en évidence dans l'intestin. Et c'est essentiellement à partir de ce réservoir qu'il va pouvoir coloniser puis infecter divers sites de l'organisme. Etudier le commensalisme de *E. coli* chez l'homme, constitue donc le *primum movens* de toute étude sur l'adaptation des clones de *E. coli* aux très nombreuses niches écologiques qui caractérisent son caractère ubiquitaire.

Le portage humain de *E. coli* est étudié depuis longtemps mais les premiers travaux étaient limités à la seule caractérisation et au seul dénombrement de l'espèce (Mitsukoa *et al.*, 1972) du fait du caractère très limité des différenciations intra spécifiques.

Depuis, les techniques moléculaires ont permis de différencier avec précision des phylums à l'intérieur de l'espèce (Escobar-Paramo *et al.*, 2004) et l'étude du commensalisme a pris en compte la détermination des groupes phylogénétiques principaux (Tenaillon *et al.*, 2010). Cependant, cette démarche restait essentiellement qualitative, basée sur le typage d'un seul clone par individu.

La validation des résultats était basée sur l'hypothèse du clone dominant qui, colonisant l'individu dès l'enfance, était statistiquement retrouvé dans ces expérimentations (Nowrouzian *et al.*, 2005).

Plus tard, certains auteurs se sont attachés à étudier plusieurs clones par individu (Schlager *et al.*, 2002, Moreno *et al.*, 2009, Lautenbach *et al.*, 2008). Mais, la faisabilité limitée de cette procédure rendait quasi impossible toute approche quantitative de la diversité intra spécifique.

Elle interdisait, en particulier, la détection précise des groupes phylogénétiques en proportions minoritaires.

D'autre part, la notion de commensalisme véritable demande une rigueur extrême dans la sélection des sujets à inclure dans une étude.

En effet, la flore intestinale est particulièrement sensible à des pressions de sélections susceptibles de la déstabiliser : antibiothérapie, hospitalisation, intervention chirurgicale, maladie intestinale dégénérative.

Elle subit également au niveau des populations humaines des variations dans le temps (Gordon *et al.*, 2005) et dans l'espace (Tenaillon *et al.*, 2010). Ainsi, la fréquence des souches intestinales du groupe B2 était en France de 5% dans les années 1980 alors qu'elle est maintenant de 30% à 40% comme dans d'autres pays industrialisés (Tableaux II et III). L'objectif de ce premier travail était donc de déterminer les proportions respectives des 4 groupes phylogénétiques principaux de *E. coli* dans les selles de sujets sélectionnés pour le caractère véritable de leur commensalisme bactérien intestinal. Pour s'affranchir des techniques phénotypiques limitantes, nous avons mis au point une technique de quantification basée sur une stratégie de PCR en temps réel.

Afin d'avoir une estimation des concentrations intestinales de *E. coli* en situation réelle de commensalisme, nous nous sommes imposés des critères rigoureux d'inclusion et d'exclusion des sujets étudiés.

Real-Time PCR for Quantitative Analysis of Human Commensal *Escherichia coli* Populations Reveals a High Frequency of Subdominant Phylogroups

Mounira Smati, Olivier Clermont, Frédéric Le Gal, Olivier Schichmanoff, Françoise Jauréguy, Alain Eddi, Erick Denamur, Bertrand Picard and for the Coliville Group
Appl. Environ. Microbiol. 2013, 79(16):5005. DOI:
10.1128/AEM.01423-13.
Published Ahead of Print 14 June 2013.

Updated information and services can be found at:
<http://aem.asm.org/content/79/16/5005>

	<i>These include:</i>
SUPPLEMENTAL MATERIAL	Supplemental material
REFERENCES	This article cites 54 articles, 33 of which can be accessed free at: http://aem.asm.org/content/79/16/5005#ref-list-1
CONTENT ALERTS	Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: <http://journals.asm.org/site/misc/reprints.xhtml>
To subscribe to to another ASM Journal go to: <http://journals.asm.org/site/subscriptions/>

Journals.ASM.org

Real-Time PCR for Quantitative Analysis of Human Commensal *Escherichia coli* Populations Reveals a High Frequency of Subdominant Phylogroups

Mounira Smati,^{a,b,c} Olivier Clermont,^{a,d} Frédéric Le Gal,^c Olivier Schichmanoff,^c Françoise Jauréguy,^{a,b,c} Alain Eddi,^e Erick Denamur,^{a,d} Bertrand Picard,^{a,b,c} for the Coliville Group

UMR-S 722, INSERM, Paris, France^a; UMR-S 722, Université Paris Nord, PRES Sorbonne Paris Cité, Bobigny, France^b; APHP, Hôpitaux Universitaires Paris Seine Saint-Denis, Site Avicenne, Bobigny, France^c; UMR-S 722, Université Paris Diderot, PRES Sorbonne Paris Cité, Paris, France^d; Département de Médecine Générale, Université Paris Diderot, PRES Sorbonne Paris Cité, Faculté de Médecine, Paris, France^e

Escherichia coli is divided into four main phylogenetic groups, which each exhibit ecological specialization. To understand the population structure of *E. coli* in its primary habitat, we directly assessed the relative proportions of these phylogroups from the stools of 100 healthy human subjects using a new real-time PCR method, which allows a large number of samples to be studied. The detection threshold for our technique was 0.1% of the *E. coli* population, i.e., 10⁵ CFU/g of feces; in other methods based on individual colony analysis, the threshold is 10%. One, two, three, or four phylogenetic groups were simultaneously found in 21%, 48%, 21%, and 8% of the subjects, respectively. Phylogroups present at a threshold of less than 10% of the population were found in 40% of the subjects, revealing high within-individual diversity. Phylogroups A and B2 were detected in 74% and 70% of the subjects, respectively; phylogroups B1 and D were detected in 36% and 32%, respectively. When phylogroup B2 was dominant, it tended not to cooccur with other phylogroups. In contrast, other phylogroups were present when phylogroup A was dominant. These data indicate a complex pattern of interactions between the members of a single species within the human gut and identify a reservoir of clones that are present at a low frequency. The presence of these minor clones could explain the fluctuation in the composition of the *E. coli* microbiota within single individuals that may be seen over time. They could also constitute reservoirs of virulent and/or resistant strains.

Escherichia coli is the most common commensal aerobic bacterium in the human gut microbiota (1); it is also the Gram-negative bacillus most frequently implicated in human extraintestinal infections (2). This apparent paradox, which characterizes opportunistic pathogens, is classically associated with host deficiencies (for example, as a result of immune-compromising illness, surgery, or catheterization). However, there are still questions concerning the role of the intrinsic properties of the parasite and the conditions that may cause commensal intestinal *E. coli* strains to become extraintestinal pathogens.

Four main phylogenetic groups named A, B1, B2, and D have been distinguished among *E. coli* strains (3, 4), and distribution within these groups appears to correlate with the origin of the strains. Phylogroup B2 has been shown to include extraintestinal virulent strains (extraintestinal pathogenic *E. coli* [ExPEC]), which express numerous virulence factors (5, 6), whereas phylogroups A and B1 contain mostly human and animal commensal strains, respectively (7, 8, 9). However, more recently, an increase in B2 phylogroup strains was observed in human commensal strains originating from industrialized countries (10, 11, 12, 13).

Little is known about the diversity, transmission, and persistence of *E. coli* commensal strains within human populations (1). An individual can be colonized by more than one distinct strain at any given time (14, 15, 16). However, the few available studies that describe strain diversity are based on the arbitrary selection of colonies from a plate inoculated with fecal samples. The typing methods used included O typing (14, 17); biotyping (17); multi-locus enzyme electrophoresis (18, 19, 20); and genotyping methods, such as pulsed-field gel electrophoresis (21), enterobacterial repetitive intergenic consensus (ERIC)-PCR (22, 23), random

amplified polymorphic DNA (RAPD) alone (24) or combined with biochemical fingerprinting (25), ribotyping (26), and triplex phylogrouping PCR (27). The studies included a small number of subjects and/or selected a relatively small number of colonies on which to base their investigations of *E. coli* population structure. The ability to accurately characterize *E. coli* strain diversity in human subjects, and the likelihood of identifying a particular *E. coli* strain, are directly related to the number of colonies sampled and the underlying prevalence of the strain (21, 26). For example, for a 90% likelihood of identifying a strain present in $\geq 10\%$ of colonies, 22 colonies must be sampled (21). New approaches are therefore needed to improve the sensitivity of detection. The detection of fecal strains present at low frequencies is necessary to understand within-individual fluctuations in strain composition that can be observed over time (18, 20, 28, 29).

To obtain insights into the population structure of *E. coli* in its primary habitat, we assessed the relative proportions of the main *E. coli* phylogenetic groups in the stools of 100 healthy human subjects; their intestinal microbiota had undergone minimal pathological and medical perturbations. A new, rapid, and sensi-

Received 1 May 2013 Accepted 7 June 2013

Published ahead of print 14 June 2013

Address correspondence to Erick Denamur, erick.denamur@inserm.fr.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AEM.01423-13>.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AEM.01423-13

TABLE 1 Primers and phylogroup-specific probes

Target organism	Primers and probe	Sequence (5'–3') ^a	Target gene	Reference or source
<i>E. coli</i>	ECFW ECRV 16S rRNA gene	CATGCCGCGTGTATGAAGAA CGGTAACGTCAATGAGCAA FAM-TATTAACCTTTACTCCCTTCCTCCCGCTGAA-TAMRA	16S rRNA	33
<i>E. coli</i> phylogroup B1	TspFW TspRV P _{TspE4.C2} B1	CAACGCTACCTTGGCGTTATC CCGCTCTCCAGGCAACATC FAM-ATCAGCGAGGCCGCGCA-TAMRA	TspE4.C2	This study
<i>E. coli</i> phylogroup B2	chuAFW chuARV P _{chuA} B2	ATTAACCCGGATACCGTTACCA CTGCTGCTGATATGTGTGAGC FAM-CGTACCGACCCAACCA-MGB	chuA	This study
<i>E. coli</i> phylogroup D	chuAFW chuARV P _{chuA} D	ATTAACCCGGATACCGTTACCA CTGCTGCTGATATGTGTGAGC FAM-CGTACCAACCCACCA-MGB	chuA	This study
<i>E. coli</i> phylogroups A ₁ and B2	yjaAFW yjaARV P _{yjaA} A ₁ /B2	CGCCTGTTAATCGCCAATTT AAAAGAATGCCAGTTGAACG FAM-AAGTTCTGCAAGATCTTGTCTGCAACTCCA-TAMRA	yjaA	This study

^a FAM, 6-carboxyfluorescein; TAMRA, 6-carboxytetramethylrhodamine.

tive real-time PCR strategy, using the same targets as triplex phylogrouping PCR (4), was applied to samples from the subjects.

MATERIALS AND METHODS

Subjects. From May 2009 to December 2011, we recruited 100 healthy human subjects from the region of Ile-de-France (Paris, France, and its suburban area). All participants lived in the community and volunteered to self-collect a fecal swab sample. The subjects had no history of gastrointestinal disease and no symptoms of immunodepression, had not received antibiotic therapy in the previous month, and had not been hospitalized in the 3 months preceding inclusion. Written informed consent was obtained from each participant, and the study was approved by the ethics evaluation committee of Institut National de la Santé et de la Recherche Médicale (INSERM) (CCTIRS no. 09.243, CNIL no. 909277, and CQI no. 01-014).

Samples and characterization of isolated strains. Fecal samples were self-collected by the subjects using Amies transport medium swabs (Venturi Transystem; Copan, Brescia, Italy) and sent by mail to the Avicenne hospital laboratory (Bobigny, France). The swabs were discharged in glycerol stock solution (Cryobank; Biovalley, Marne la Vallée, France) and stored at –80°C until they were used. The quantity of feces in 0.2 ml of solution was estimated in preliminary experiments to be 10 mg; *E. coli* counts from 10 swabs were compared to those from corresponding fresh feces, which were chosen for a variety of consistencies (data not shown). The stool-containing suspensions were plated onto Drigalski agar plates (Bio-Rad, Life Science, Marnes-la-Coquette, France). After 24 h of incubation at 37°C, one colony was randomly picked, identified as *E. coli* using API 20E (bioMérieux, Marcy l'Etoile, France), and stored in glycerol stock solution. Twenty to 30 *E. coli* colonies were randomly picked from each of 20 original plates and stored in glycerol stock solution at –80°C to serve as controls. The phylogenetic groups (A, B1, B2, and D) of the randomly picked *E. coli* strains were determined by the triplex PCR method (4) and by the recently reported quadriplex method, which also detects *Escherichia* clade strains (13, 30). The strains were tested for their antibiotic susceptibilities using the disk diffusion method according to the 2012 recommendations of the Comité de l'Antibiogramme de la Société Française de Microbiologie (Antimicrobial Committee of the French Society for Microbiology; <http://www.sfm-microbiologie.org>). The following antimicrobial agents were tested: amoxicillin, amoxicillin-clavulanic acid, cefoxitin, ceftriaxone, amikacin, ofloxacin, and sulfamethoxazole-trimethoprim.

DNA extraction from feces. The QIAamp DNA Stool Minikit (Qia-gen, Courtabœuf, France) was used to extract DNA from 0.2 ml of frozen stool sample according to the manufacturer's recommendations with modifications. An internal DNA control corresponding to the R' 1 region of the hepatitis delta virus (HDV) genome prepared as previously described (31) was added, before extraction, at 10⁵ copies per sample. The DNA was eluted in a final volume of 200 µl and stored at –80°C.

Quantitative-PCR assay. (i) Design of primers and phylogroup-specific probes. The sequences used in the triplex PCR phylogrouping method (4) (TspE4.C2, *chuA*, and *yjaA*) were obtained from GenBank, EMBL, and the Broad Institute (<http://www.broadinstitute.org>) databases (from the whole-genome sequences of *Escherichia* strains) and aligned using the Clustal W program (provided by the European Bioinformatics Institute). The strains were assigned *in vitro* or *in silico* to their corresponding phylogroups by the triplex PCR method and/or by multilocus sequence typing (MLST). Primers were designed to anneal to conserved sequences, whereas probes were designed to target unique sites, allowing specific detection of each phylogroup with the Primer Express 3.0 software (Applied Biosystems, Villebon Sur Yvette, France) and the Primer3 Plus online interface (32) (provided by the Whitehead Institute). A list of strains tested *in silico* is given in Table S1 in the supplemental material. Two strains belonging to phylogroup D (TA255 and TA280) have a phylogroup F *chuA* sequence (probably resulting from horizontal gene transfer) and were identified as B2 by our method. Primers and probes were manufactured by Sigma-Aldrich (Lyon, France) and Applied Biosystems, respectively. All primers and probes used are listed in Table 1.

(ii) Phylogroup-specific real-time PCR. Purified DNA (20 µl) was added to 30 µl of PCR mixture containing 25 µl of TaqMan Universal PCR master mix II (2×; Applied Biosystems), 300 nM each primer, 100 nM fluorescent probe, and bovine serum albumin at a final concentration of 0.1 µg/µl (New England BioLabs, Evry, France). However, for the B1 phylogenetic group, primers (TspFW and TspRV) and probe (P_{TspE4.C2}B1) were used at 900 nM and 250 nM, respectively. The reaction mixture was heated to 50°C for 2 min (initiation step) and then 95°C for 10 min. This was followed by 45 cycles of amplification that each consisted of 15 s at 95°C and 1 min at 60°C. A no-template control and a positive-control sample, which had been quantified previously, were included in each run. The run was considered valid if the interassay coefficient of variation (CV) for the control sample was less than 30%. Each assay was performed in duplicate in the same run. The DNA extraction was consid-

ered to be valid if the difference between the cycle threshold (C_T) of the internal-control HDV and the average C_T (calculated as the average of the HDV C_T values for all samples in the series) was less than the standard deviation. The reactions, data acquisition, and analyses were performed using the ABI Prism 7000 sequence detection system (Applied Biosystems). Standard curves made from known concentrations of DNA for each set of primers and probes were used for quantification.

(iii) Preparation of phylogroup-specific PCR standards. Quantification controls were obtained using one representative strain for each phylogroup: K-12 (phylogroup A), IAI1 (phylogroup B1), ED1a (phylogroup B2), and UMN026 (phylogroup D). The Genomic DNA Buffer Set (Qiagen) and Qiagen genomic tips 100 G (Qiagen) were used to extract DNA from bacterial cultures according to the manufacturer's recommendations. Bacterial counts were obtained by plating, the concentration of purified DNA was determined with a spectrophotometer, and the corresponding copy number was calculated according to genome length. The two methods of quantification gave similar results (data not shown). Serial 10-fold dilution series for 8×10^7 to 12×10^7 down to 0.8×10^1 to 1.2×10^1 target genomes were applied for real-time PCR.

(iv) *E. coli* real-time PCR quantification assay. Quantitative PCR for *E. coli* 16S rRNA genes was performed as previously described (33) using 20 μ l of DNA purified from stool samples.

Control strains. The following bacteria were used to evaluate the specificity of PCR primers and probes: *E. coli sensu stricto* ED1a, CFT073, 536, S88, IAI1, 55989, IAI39, ECOR36, DAECT14, K-12 MG1655, HS, H10407, 042, and UMN026; *Escherichia* clade I (ROAR 185 and M863), clade II (ROAR 19), clade III (ROAR 291 and ROAR 438), clade IV (E243 and B49), and clade V (ROAR 292 and ROAR 129); *Escherichia fergusonii* ATCC 35469; *Escherichia albertii* CIP107988; *Escherichia blattae* CIP104942; *Escherichia vulneris* CIP103177; *Escherichia hermannii* CIP103176; and *Salmonella enterica* serovar Typhimurium, *Klebsiella pneumoniae*, *Proteus mirabilis*, *Pseudomonas aeruginosa* ATCC 27853, *Enterococcus faecalis* ATCC 33186, *Enterococcus faecium*, *Streptococcus bovis*, *Bacteroides fragilis*, *Bacteroides thetaiotaomicron*, *Bacteroides vulgatus*, *Fusobacterium nucleatum*, *Prevotella melaninogenica*, *Clostridium difficile*, *Clostridium perfringens*, *Bifidobacterium adolescentis*, *Bifidobacterium breve*, *Eubacterium exiguum*, *Eubacterium lentum*, *Lactobacillus casei*, and *Lactobacillus cateniformis*.

Statistical analyses. A factorial analysis of correspondence (FAC) (34) was used to describe associations among the data; a two-way table was analyzed using SPAD.N software (Cisia, Saint Mandé, France). The table had 98 rows (corresponding to the 98 subjects with *E. coli* identified in their stool samples) and 22 columns (corresponding to the 10 variables). The 10 variables were the phylogenetic group (A, B1, B2, or D) of the unique randomly selected strain, the dominant phylogenetic group, the intermediate phylogenetic group, and the minor phylogenetic group (defined below); the presence of a high quantitative level of *E. coli* in stools (more than 10^8 CFU per gram); subject gender; subject age (above or below 60 years old); and strain resistance to amoxicillin, ofloxacin, or sulfamethoxazole-trimethoprim. We used a binary code for each variable: present, 1; absent, 0. The FAC uses a covariance matrix based on χ^2 distances. The computation determines a plane defined by two principal axes of the analysis. The first axis, F1, accounts for most of the variance, and the second axis, F2 (orthogonal to F1), accounts for the largest part of the variance not accounted for by F1 (34). The variables of the unique randomly selected strains were used as illustrative variables. All other variables were used to compute the plane.

RESULTS

Characteristics of the study population and of randomly selected *E. coli* commensal strains (one per individual). The study included 46 males and 54 females between the ages of 26 and 86 years (median, 58.0 ± 11.7 years). All but two of the subjects carried *E. coli*. We studied one randomly selected clone per individual, which was considered to be the dominant clone. The resis-

TABLE 2 *E. coli* phylogenetic group and subgroup determination using the results of PCR amplification of the *chuA* gene, *yjaA* gene, and DNA fragment TspE4.C2^a

<i>E. coli</i> phylogenetic group and subgroup	Presence of the sequence ^b		
	<i>chuA</i>	<i>yjaA</i>	TspE4.C2
A ₀	—	—	—
A ₁	—	+	—
B1	—	—	+
B2 ₂	+	+	—
B2 ₃	+	+	+
D ₁	+	—	—
D ₂	+	—	+

^a As in reference 4.

^b +, present; —, absent.

tance levels of these 98 *E. coli* dominant clones to amoxicillin, sulfamethoxazole-trimethoprim, amoxicillin-clavulanic acid, and ofloxacin were 25%, 14%, 7%, and 2%, respectively. No resistance to broad-spectrum cephalosporins or amikacin was detected.

The classical phylogroup triplex PCR method used on the dominant clones showed the prevalence of the different phylogroups to be as follows: phylogroup A, 31%; phylogroup B1, 13%; phylogroup B2, 33%; and phylogroup D, 21%.

Stool sample quantification of the four main *E. coli* phylogroups. We developed a real-time PCR assay targeting the genes involved in the classical phylogrouping triplex PCR method (4) to directly quantify the four main *E. coli* phylogroups (A, B1, B2, and D) from stool samples. In this assay, the phylogroups B1 (*chuA* negative, *yjaA* negative, and TspE4.C2 positive), B2 (*chuA* positive, *yjaA* positive, and TspE4.C2 variable), and D (*chuA* positive, *yjaA* negative, and TspE4.C2 variable) were quantified by real-time PCR using probes specifically targeting B1 TspE4.C2, B2 *chuA*, and D *chuA* sequences, respectively (Table 2). The proportion of each phylogroup was calculated as a percentage of the total *E. coli* count obtained by quantitative PCR for 16S rRNA genes (33). The proportion of phylogroup A, which encompasses subgroups A₀ (*chuA*, *yjaA*, and TspE4.C2 negative) and A₁ (*chuA* negative, *yjaA* positive, and TspE4.C2 negative) (Table 2), was calculated by subtracting the proportions of phylogroups B1, B2, and D from the total *E. coli* count. The proportion of subgroup A₁ was confirmed using an *yjaA*-specific probe ($p_{yjaA A_1/B2}$), which quantifies the A₁ subgroup and the B2 phylogroup using the following formula: $A_1 = yjaA^{\text{positive}} - B2$. This $p_{yjaA A_1/B2}$ probe was used to estimate the lower limit of detection of the proportion of phylogroup A. The comparison between the proportion of phylogroup A (calculated by subtraction) and the proportion of subgroup A₁ (quantified by the $p_{yjaA A_1/B2}$ probe) allowed us to estimate the threshold of phylogroup A detection by subtraction. Concordance was observed when the proportion of phylogroup A was at least 15% of the total *E. coli* count (data not shown). With this approach and using four probes (16S rRNA gene, $p_{TspE4.C2 B1}$, $p_{chuA B2}$, and $p_{chuA D}$), we estimated that phylogroup A was actually present at over 15% and that real-time PCR variations did not make a significant contribution to the results. This threshold value was further supported by comparison to PCR findings for several clones of plated fecal samples (see below).

Specificity, sensitivity, and reproducibility of the assay. (i) Assay specificity. The specificity of the *E. coli* primers and probes was confirmed using an extensive set of control strains; this set

(mainly gastrointestinal tract bacteria of non-*Escherichia* species [see Materials and Methods]) produced negative PCR results. The specificity of the 16S rRNA gene assay (33), as well as that of the phylogroup probes and primers, was tested using DNAs from various phylogenetic groups of *E. coli sensu stricto* and from closely related bacterial species of the genus *Escherichia* (including the recently described *Escherichia* clades [30, 35]). The 16S rRNA gene probe detects the phylogroups of *E. coli sensu stricto*, *E. albertii*, *E. fergusonii*, and the *Escherichia* clades, but not *E. blattae*, *E. hermannii*, and *E. vulneris*, in accordance with their taxonomic status (36, 37).

Each phylogroup-specific probe gave positive PCR results for the corresponding target bacteria and negative PCR results for nontarget microorganisms, with a small number of exceptions. The *Escherichia* clade III and IV strains, and some strains of clade I (ROAR 185), were recognized by the *chuA* B2 probe. The strains of the recently reported phylogroup F, defined by MLST (38), are considered D₁ subgroup by the triplex PCR method (Table 2) but were split into B2 and D phylogroups by the real-time PCR method. This split corresponds to the different branches of the phylogroup F strains on the MLST tree (39). We did not detect a signal above the threshold for the nontemplate controls in any of the assays ($C_T > 40$).

Spiking experiments were carried out to test the ability of our assay to pick out a specific bacterial DNA from a complex DNA background. Defined amounts of DNA from each of the studied phylogenetic groups B1, B2, and D were quantified by real-time PCR and spiked with DNA belonging to other phylogenetic group strains. No significant difference was observed, indicating high specificity of the probes used and no evidence of cross-reactivity.

(ii) Assay sensitivity, linearity, and efficiency. The detection limits of the real-time PCR assays were examined using 10-fold serial dilutions of target DNAs extracted from pure culture of *E. coli* A, B1, B2, and D. The sensitivity of the probe specific for $p_{yjaA}A_1/B_2$, $p_{chuA}B_2$, $p_{chuA}D$, and $p_{TspE4.C2}B1$ was below 10 CFU per PCR, which corresponded to approximately 9×10^4 CFU/g feces. The cycle number at which product fluorescence exceeded a defined threshold varied linearly over a range of concentrations from 10^6 to 10 bacteria per PCR mixture ($r^2 = 0.99$) (data not shown). The linear range of all assays was therefore 9×10^9 to 9×10^4 bacteria per g of feces. The efficiencies of the PCRs were around 95%, as the slopes of the standard curves were close to the optimal theoretical values (see Table S2 in the supplemental material).

(iii) Assay reproducibility. Based on the C_T values of 15 replicates, the intra-assay reproducibility was found to be high (CV of C_T values, <1%, corresponding to a CV of copies of <20%) (see Table S2 in the supplemental material).

Comparison with PCR typing of individual clones. To test the robustness of our real-time PCR results, we compared our phylogroup quantification to the triplex PCR typing of 20 to 30 clones from 20 subjects. From the number of randomly selected colonies, the probability of detecting a minor clone (defined as a clone constituting up to 10% of the *E. coli* population) is estimated to be >90% (20). We found a correlation between our quantitative-PCR assay and the phylogroup determination of individual clones. In all cases, the phylogroups detected by the classical PCR typing of individual clones were also detected by the quantitative-PCR assay; we found similar quantitative trends

TABLE 3 Distribution of phylogenetic groups according to their proportions determined by quantitative PCR of fecal samples from 100 healthy volunteers

Phylogroup	Distribution (%) ^a		
	DPG	IPG	MPG
A	38	27	35
B1	13	10	77
B2	37	13	50
D	11	5	84

^a DPG, dominant phylogenetic group (>50%); IPG, intermediate phylogenetic group (10 to 50%); MPG, minor phylogenetic group (<10%) (20).

but some differences in the proportions of the phylogroups. Interestingly, the real-time PCR assay detected minor clones, not found by plating, in 40% of the cases (see Table S3 in the supplemental material).

Quantitative determination of *E. coli* phylogenetic group composition. Preliminary global quantification of *E. coli* in the stools of the subjects was carried out using real-time PCR with the 16S rRNA gene *E. coli* probe and primers (33). Negative real-time PCR results were found for the two subjects for whom *E. coli* also failed to be detected by plating. Excluding these two subjects, we found an average of 7.84 ± 0.54 log CFU of *E. coli* per gram of stool. The respective proportions of the four phylogenetic groups A, B1, B2, and D were determined for the 98 subjects for whom *E. coli* was detected. To verify that the use of transportation swabs sent by regular mail did not bias *E. coli* quantification, we compared real-time PCR results obtained from eight fresh control fecal samples and from their corresponding swabs, which were collected using the same procedure described in Materials and Methods, and no significant difference was found (see Table S4 in the supplemental material), in accordance with recent data on storage conditions of samples (40). However, if possible, analysis of fresh feces is preferred.

The phylogenetic groups A, B1, B2, and D were detected in 74%, 36%, 70%, and 32% of the subjects, respectively; the groups A and B2 were more frequently detected than B1 and D. However, the proportions of the four phylogroups varied widely among the subjects. The within-subject prevalence of each phylogenetic group varied from 0 to 100%. One, two, three, or four phylogenetic groups were found in 21%, 48%, 21%, and 8% of the subjects, respectively. When only a single group was detected, it was most commonly phylogroup B2 (62%); phylogroup A was found in 19% of subjects and phylogroups B1 and D in 9.5% of subjects. When the proportions of the phylogroups were defined according to three categories, as suggested by Schlager et al. (20) (dominant phylogenetic group, >50% of the *E. coli* population; intermediate phylogenetic group, 10 to 50% of the *E. coli* population; and minor phylogenetic group, <10% of the *E. coli* population), the proportions of the three categories varied significantly according to the four phylogenetic groups (Table 3). Phylogroups B2 and A were equally prevalent as the dominant group and in a higher proportion of samples than phylogroup B1 or D. Phylogroup A was the most common intermediate phylogenetic group, and phylogroup D was the most common minor phylogenetic group. Minor phylogroups (less than 10% of the total *E. coli* population) were detected in 40% of the samples.

When the 12 combinations (three categories and four phylogenetic groups) were considered, the proportions of the cat-

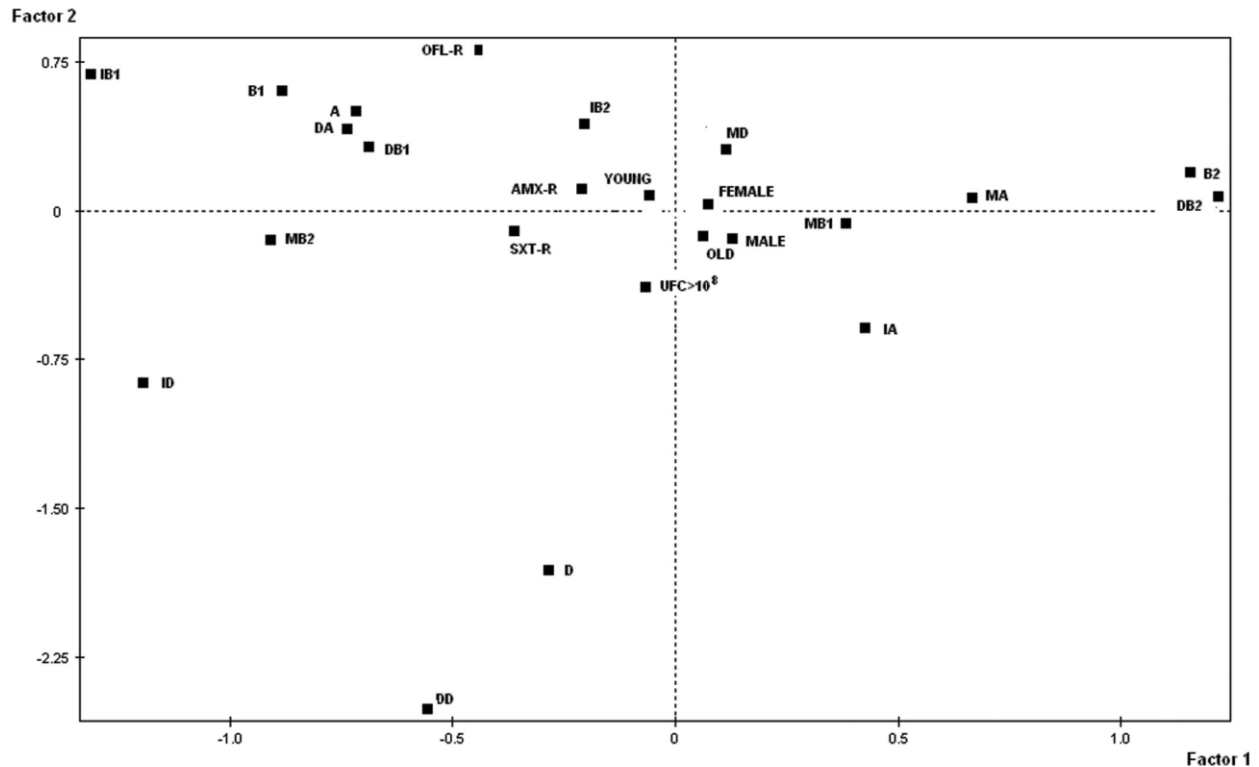


FIG 1 FAC for the 98 subjects with *E. coli* found in their feces. Projections on the plane F1-F2 of the four phylogenetic groups (A, B1, B2, and D) of the unique randomly selected strains, of the dominant phylogenetic group (DA, DB1, DB2, or DD), of the intermediate phylogenetic group (IA, IB1, IB2, or ID), and of the minor phylogenetic group (MA, MB1, MB2, or MD) defined for each of the four phylogenetic groups are shown. The dominant, intermediate, and minor groups were defined from data from the quantitative-PCR assay. The abundance of *E. coli* in stools (CFU $> 10^8$); the gender of the subjects (male or female); the age of the subjects (older or younger than 60 years); and the resistance of the unique randomly selected strains to amoxicillin (AMX-R), ofloxacin (OFL-R), and sulfamethoxazole-trimethoprim (SXT-R) were also projected on the plane.

egories varied according to the phylogenetic group considered. This complex variation is described by a FAC (Fig. 1). On the F1-F2 plane of the FAC, which accounted for 32.7% of the total variance, the dominant phylogenetic group B2 was distinguished by the positive values of the first factor, whereas the dominant phylogenetic groups B1, A, and D were distinguished by the negative values of this factor. The intermediate phylogenetic groups B1, B2, and D were projected on the negative values of the F1 axis and thus appeared to be associated with the dominant phylogenetic groups A and B1. The intermediate phylogenetic group A and the minor phylogenetic groups A, B1, and D were projected on the positive values of the F1 axis and thus are associated with the dominant phylogenetic group B2. This analysis allows us to distinguish between two types of structures within the *E. coli* intestinal commensal populations, according to the major dominant group B2 or A. The dominant phylogenetic group B2 was associated with the absence of the other groups, whereas the dominant phylogenetic group A was associated with variable proportions of other groups. The second axis, F2, characterized the dominant phylogenetic group D by its negative value. No link with the absolute quantity of *E. coli* bacteria was observed, as the variable $> 10^8$ CFU was projected near the origin of the axes and thus was not related to phylogenetic repartition.

***E. coli* phylogenetic group composition: relevance of subject characteristics and the phylogroup and antibiotic resistance of a randomly selected clone.** There is no significant difference in the phylogenetic group composition according to the age or gender of the subjects: these variables were projected near the origin of the axes on the FAC (Fig. 1).

In 78% of the cases, the dominant phylogenetic group determined by the real-time PCR was identical to the one determined by the method of Clermont et al. (4) on one randomly selected clone. In 22% of the cases, the phylogenetic group of the randomly selected clone was present in the sample, but as an intermediate or minor phylogenetic group. There was a discrepancy in six cases, which corresponded to the selected clone being identified as phylogroup D whereas the quantitative PCR identified it as belonging to phylogroup B2. We have shown that in all six of these cases, the randomly selected clone belonged to phylogroup F (13): this is supported by the fact that the $p_{chuA}B2$ probe that we designed is positive for some F group strains (see Table S3 in the supplemental material). The good correlation between the quantitative-PCR assay and classical phylogroup determination (based on one randomly selected clone) is shown in Table 3, where the proportions of the dominant phylogenetic groups (A, 38%; B1, 13%; B2, 37%; D, 11%) are close to the proportions detected in the randomly selected clones (A, 31%; B1, 13%; B2, 33%; D, 21%), as well as on

the FAC (Fig. 1). The phylogenetic group of the randomly selected clone (considered an illustrative variable) is projected close to the dominant phylogenetic group on the plane, with the exception of the D phylogroup variable on the negative values of F2, which is attracted by the dominant phylogroup B2 due to the six cases of discrepancy. The resistance levels of the randomly selected clones to amoxicillin, sulfamethoxazole-trimethoprim, or ofloxacin were also considered illustrative variables. They were associated with the dominant A phylogroup for negative values of F1 and distantly related to the dominant B2 phylogroup on the FAC (Fig. 1).

DISCUSSION

We have developed an original quantitative-PCR assay, based on the classical phylotyping method of Clermont and colleagues (4), to determine the relative proportions of the four main *E. coli* phylogroups, A, B1, B2, and D. On average, 7.84 ± 0.54 log CFU of *E. coli* were found per gram of feces, in accordance with previous studies (33, 41, 42). When assessing the population structure of intestinal commensal bacteria in humans, it is necessary to study subjects whose microbiotas are not perturbed by external factors, such as antibiotic consumption, hospitalization, or intestinal disease. We were able to study real commensal *E. coli* populations, as our 100 healthy subjects conformed to strict inclusion criteria; we excluded any potential subject with a physiopathological condition that could disturb the intestinal microbiota equilibrium. The level of acquired resistance to antibiotics in our study population was much lower than that reported in community-acquired *E. coli* urinary tract infections in France (ONERBA 2008 and AFORCOPI-BIO network 2007 [<http://www.onerba.org>]): amoxicillin, 25% versus 44%; sulfamethoxazole-trimethoprim, 14% versus 20%; amoxicillin-clavulanic acid, 7% versus 24%; and ofloxacin, 2% versus 11%. This supports our belief that the study populations are commensal and in the same range of resistance found in a recent study in Paris on fecal strains (43).

The specificity of our assay is satisfactory, although it has some limitations. Some strains of the recently reported *Escherichia* clades (30, 35) gave a signal with the p_{chuA} B2 probe, which can falsely increase the proportion of phylogroup B2. However, as these clade strains were *yjaA* negative, the opposite of the *E. coli* B2 strains, a strong B2 signal with a weak or negative *yjaA* signal should indicate the presence of clade I, III, or IV strains. The prevalence of these *Escherichia* clades in human commensal populations is very low (around 2 to 3%) (35) and thus will not significantly affect the results of our assay. We did not detect any *Escherichia* clade among the 100 randomly selected clones using the new Clermont phylotyping method (13) (see Table S3 in the supplemental material). Some strains classified as phylogroup D by the classical triplex PCR method and belonging to the recently reported phylogroup F (identified by MLST [38]) appeared to belong to phylogroup B2 due to a positive signal with the p_{chuA} B2 probe. Phylogroup F is closely related to phylogroup B2 (1). This probe cross-reactivity was responsible for all six discrepancies observed between our quantitative-PCR assay and the classical triplex PCR phylogrouping of the randomly selected clones (see Table S3 in the supplemental material). However, it may not significantly alter the results, as phylogroup F strains have been found to represent less than 8% of the population in a large cohort of fecal strains from Australia and France (13).

We have shown that our assay is highly reproducible, with a detection threshold of 10^5 CFU/g of feces; this corresponds to

0.1% of the *E. coli* population, much lower than the 10% threshold obtained by other methods (20, 21, 23, 26, 29). Previous studies picked 20 to 30 isolates from an agar plate per individual and used a semiquantitative method of assessment (20, 23, 25). We found good correlation between data obtained with such an approach and that obtained using our quantitative method (see Table S3 in the supplemental material). However, real-time PCR may be faster, easier to perform, and more suitable for large sample sizes. In 40% of cases (8/20), real-time PCR detected phylogroups that were not found by plating. Our specific probes were able to detect groups with very low prevalence thanks to a low detection threshold. To reach a level of detection (0.1%) similar that in our assay by the alternative semiquantitative method, a large number of isolates would have to be studied from each fecal sample. The classical qualitative method appeared to underestimate the presence of minor groups. This explains the apparent differences in fecal carriage of the *E. coli* phylogroups (e.g., the prevalence of the B2 phylogroup in the studied population is apparently 33% if only major clones are considered but 70% if all clones, including minor clones, are considered). We acknowledge that, in giving a global estimation of the structure of the commensal strains, our method does not allow us to study the *E. coli* clones separately.

Each person commonly carries a dominant strain of *E. coli* that constitutes more than half of the total colonies isolated (other strains are present at various levels) (14, 15, 16, 18, 44). Depending on the study, between 92% and 100% of human samples contained a dominant clone representing more than 50% of the *E. coli* population (17, 20, 21, 23). When we compare the results of the quantitative PCR with those based on a qualitative approach in which a single clone was randomly selected per individual, we obtained 78% correlation if the prevalence of a given phylogenetic group is 50% (see Table S3 in the supplemental material). Our method ensures correct identification of *E. coli* dominant-strain phylogroups.

The prevalence of *E. coli* phylogroups in the stools of humans has been shown to depend on sex, age, year of sampling, country of origin, and diet (1, 12, 45). In our population and using our approach, phylogroups A and B2 were detected in 74% and 70% of cases, respectively, whereas phylogroups B1 and D were detected less often. Phylogroups A and B2 were overrepresented among the dominant phylogroups (Table 3). These results are in accordance with previous data (reviewed in reference 1) and confirm the prominent role played now by phylogroup B2, as well as phylogroup A, in commensal microbiotas of French subjects. The B2 phylogroup strains appear to be opposed to the B1 phylogroup strains, which are more animal specific (9). We found no correlations to be associated with gender or age in our samples.

High intrahost diversity was found, as two, three, and four phylogenetic groups were simultaneously detected in 48%, 21%, and 8% of subjects, respectively. There was a nonrandom, complex level of phylogroup combinations within single subjects. As illustrated in the FAC analysis (Fig. 1), there is antagonism between phylogroups B1 and B2. This could correspond to different ecological niches defined, for example, by different diets. Among the most prevalent phylogroups (A and B2), there appear to be distinct population structures. A dominant phylogroup B2 was associated with an absence of other groups, whereas the presence of the other phylogroups was detected with a dominant phylogroup A (Fig. 1). These data support the notion that extraintestinal virulence may be a by-product of commensalism: numerous

“extraintestinal virulence genes” (coding for adhesins, iron capture systems, toxins, and protectins) exhibited by phylogroup B2 strains (6, 46) may have evolved primarily to allow the bacteria to be good intestinal colonizers (24, 28, 47–49). When phylogroup B2 is dominant, it does not cooccur with the residence of other phylogroup strains, possibly due to higher fitness (50, 51). Alternatively, dominance of phylogroup B2 may reflect the presence of a specific gut environment or parasitism by viruses or bacteriocins that favor it. As previously reported (52), an association between the non-B2 phylogroup strains, especially phylogroup A, is observed with antibiotic resistance (Fig. 1). This is of particular interest, as the intestinal microbiota plays a critical role in the emergence of antibiotic resistance (53, 54, 55). Further work is needed to investigate the characteristics of the hosts corresponding to the two patterns of dominant phylogroups. This would help us to identify the ecological forces shaping population structure and to better understand the impact of exposure to opportunistic *E. coli* extraintestinal infections and the emergence of antibiotic resistance.

We have developed a rapid and reliable real-time PCR assay that quantifies *E. coli* phylogroups directly from stool samples and detects, with high frequency, the presence of subdominant clones. The presence of minor clones could explain the fluctuation in the composition of the *E. coli* microbiota within single individuals that may be seen over time. They could also constitute reservoirs of virulent and/or resistant strains. Long-term studies of subject cohorts, using our approach, will allow these notions to be tested.

ACKNOWLEDGMENTS

We are grateful to Hervé Jacquier for providing some of the control strains.

The Coliville Group is composed of the following medical practitioners who were involved in the recruitment of subjects for the study: Monique Allouche, Jean-Pierre Aubert, Isabelle Aubin, Ghislaine Audran, Dan Baruch, Philippe Birembaux, Max Budowski, Emilie Chemla, Alain Eddi, Marc Frasier, Eric Galam, Julien Gelly, Serge Joly, Jean-François Millet, Michel Nougairède, Nadja Pilon, Guy Septavaux, Catherine Szewbel, Philippe Vellard, Raymond Wakim, Xavier Watelet, and Philippe Zerr.

REFERENCES

- Tenaillon O, Skurnik D, Picard B, Denamur E. 2010. The population genetics of commensal *Escherichia coli*. *Nat. Rev. Microbiol.* 8:207–217.
- Russo TA, Johnson JR. 2003. Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes Infect.* 5:449–456.
- Herzer PJ, Inouye S, Inouye M, Whittam TS. 1990. Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. *J. Bacteriol.* 172:6175–6181.
- Clermont O, Bonacorsi S, Bingen E. 2000. Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl. Environ. Microbiol.* 66:4555–4558.
- Bingen E, Picard B, Brahimi N, Mathy S, Desjardins P, Elion J, Denamur E. 1998. Phylogenetic analysis of *Escherichia coli* strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent B2 group strains. *J. Infect. Dis.* 177:642–650.
- Picard B, Garcia JS, Gouriou S, Duriez P, Brahimi N, Bingen E, Elion J, Denamur E. 1999. The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect. Immun.* 67:546–553.
- Duriez P, Clermont O, Bonacorsi S, Bingen E, Chaventré A, Elion J, Picard B, Denamur E. 2001. Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. *Microbiology* 147:1671–1676.
- Gordon DM, Cowling A. 2003. The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology* 149:3575–3586.
- Escobar-Páramo P, Le Menac’h A, Le Gall T, Amorin C, Gouriou S, Picard B, Skurnik D, Denamur E. 2006. Identification of forces shaping the commensal *Escherichia coli* genetic structure by comparing animal and human isolates. *Environ. Microbiol.* 8:1975–1984.
- Obata-Yasuoka M, Ba-Thein W, Tsukamoto T, Yoshikawa H, Hayashi H. 2002. Vaginal *Escherichia coli* share common virulence factor profiles, serotypes and phylogeny with other extraintestinal *E. coli*. *Microbiology* 148:2745–2752.
- Zhang L, Foxman B, Marrs C. 2002. Both urinary and rectal *Escherichia coli* isolates are dominated by strains of phylogenetic group B2. *J. Clin. Microbiol.* 40:3951–3955.
- Gordon DM, Stern SE, Collignon PJ. 2005. Influence of the age and sex of human hosts on the distribution of *Escherichia coli* ECOR groups and virulence traits. *Microbiology* 151:15–23.
- Clermont O, Christenson JK, Denamur E, Gordon DM. 2013. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ. Microbiol. Rep.* 5:58–65.
- Sears HJ, Brownlee I, Uchiyama JK. 1950. Persistence of individual strains of *Escherichia coli* in the intestinal tract of man. *J. Bacteriol.* 59:293–301.
- Sears HJ, Brownlee I. 1952. Further observations on the persistence of individual strains of *Escherichia coli* in the intestinal tract of man. *J. Bacteriol.* 63:47–57.
- Sears HJ, Janes H, Saloum R, Brownlee I, Lamoreaux LF. 1956. Persistence of individual strains of *Escherichia coli* in man and dog under varying conditions. *J. Bacteriol.* 71:370–372.
- Lidin-Janson G, Kaijser B, Lincoln K, Olling S, Wedel H. 1978. The homogeneity of the faecal coliform flora of normal school-girls, characterized by serological and biochemical properties. *Med. Microbiol. Immunol.* 164:247–253.
- Caugant DA, Levin BR, Selander RK. 1981. Genetic diversity and temporal variation in the *E. coli* population of a human host. *Genetics* 98:467–490.
- Caugant DA, Levin BR, Selander RK. 1984. Distribution of multilocus genotypes of *Escherichia coli* within and between host families. *J. Hyg. (London)* 92:377–384.
- Schlager TA, Hendley JO, Bell AL, Whittam TS. 2002. Clonal diversity of *Escherichia coli* colonizing stools and urinary tracts of young girls. *Infect. Immun.* 70:1225–1229.
- Lautenbach E, Bilker WB, Tolomeo P, Maslow JN. 2008. Impact of diversity of colonizing strains on strategies for sampling *Escherichia coli* from fecal specimens. *J. Clin. Microbiol.* 46:3094–3096.
- Moreno E, Andreu A, Pigrau C, Kuskowski MA, Johnson JR, Prats G. 2008. Relationship between *Escherichia coli* strains causing acute cystitis in women and the fecal *E. coli* population of the host. *J. Clin. Microbiol.* 46:2529–2534.
- Moreno E, Johnson JR, Pérez T, Prats G, Kuskowski MA, Andreu A. 2009. Structure and urovirulence characteristics of the fecal *Escherichia coli* population among healthy women. *Microbes Infect.* 11:274–280.
- Nowrouzian FL, Wold AE, Adlerberth I. 2005. *Escherichia coli* strains belonging to phylogenetic group B2 have superior capacity to persist in the intestinal microflora of infants. *J. Infect. Dis.* 191:1078–1083.
- Vollmerhausen TL, Ramos NL, Gündogdu A, Robinson W, Brauner A, Katouli M. 2011. Population structure and uropathogenic virulence-associated genes of faecal *Escherichia coli* from healthy young and elderly adults. *J. Med. Microbiol.* 60:574–581.
- Anderson MA, Whitlock JE, Harwood VJ. 2006. Diversity and distribution of *Escherichia coli* genotypes and antibiotic resistance phenotypes in feces of humans, cattle, and horses. *Appl. Environ. Microbiol.* 72:6914–6922.
- Escobar-Páramo P, Grenet K, Le Menac’h A, Rode L, Salgado E, Amorin C, Gouriou S, Picard B, Rahimy MC, Andrement A, Denamur E, Ruimy R. 2004. Large-scale population structure of human commensal *Escherichia coli* isolates. *Appl. Environ. Microbiol.* 70:5698–5700.
- Nowrouzian FL, Adlerberth I, Wold AE. 2006. Enhanced persistence in the colonic microbiota of *Escherichia coli* strains belonging to phylogenetic group B2: role of virulence factors and adherence to colonic cells. *Microbes Infect.* 8:834–840.
- De Muinck EJ, Oslashien T, Storø O, Johnsen R, Stenseth NC, Ronningen KS, Rudi K. 2011. Diversity, transmission and persistence of

- Escherichia coli* in a cohort of mothers and their infants. Environ. Microbiol. Rep. 3:352–359.
30. Walk ST, Alm EW, Gordon DM, Ram JL, Toranzos GA, Tiedje JM, Whittam TS. 2009. Cryptic lineages of the genus *Escherichia*. Appl. Environ. Microbiol. 75:6534–6544.
 31. Le Gal F, Gordien E, Affolabi D, Hanslik T, Alloui C, Dény P, Gault E. 2005. Quantification of hepatitis delta virus RNA in serum by consensus real-time PCR indicates different patterns of virological response to interferon therapy in chronically infected patients. J. Clin. Microbiol. 43:2363–2369.
 32. Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM. 2007. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res. 35:W71–W74.
 33. Penders J, Thijs C, Vink C, Stelma FF, Snijders B, Kummeling I, van den Brandt PA, Stobberingh EE. 2006. Factors influencing the composition of the intestinal microbiota in early infancy. Pediatrics 118:511–521.
 34. Greenacre M. 1992. Correspondence analysis in medical research. Stat. Methods Med. Res. 1:97–117.
 35. Clermont O, Gordon DM, Brisse S, Walk ST, Denamur E. 2011. Characterization of the cryptic *Escherichia* lineages: rapid identification and prevalence. Environ. Microbiol. 13:2468–2477.
 36. Hartl DL. 1992. Population genetics of microbial organisms. Curr. Opin. Genet. Dev. 2:937–942.
 37. Priest FG, Barker M. 2010. Gram-negative bacteria associated with brewery yeasts: reclassification of *Obesumbacterium proteus* biogroup 2 as *Shimwellia pseudoproteus* gen. nov., sp. nov., and transfer of *Escherichia blattae* to *Shimwellia blattae* comb. nov. Int. J. Syst. Evol. Microbiol. 60: 828–833.
 38. Jauregui F, Landraud L, Passet V, Diancourt L, Frapy E, Guigon G, Carbonnelle E, Lortholary O, Clermont O, Denamur E, Picard B, Nassif X, Brisse S. 2008. Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains. BMC Genomics 9:560.
 39. Clermont O, Olier M, Hoede C, Diancourt L, Brisse S, Keroudean M, Glodt J, Picard B, Oswald E, Denamur E. 2011. Animal and human pathogenic *Escherichia coli* strains share common genetic backgrounds. Infect. Genet. Evol. 11:654–662.
 40. Lauber CL, Zhou N, Gordon JI, Knight R, Fierer N. 2010. Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. FEMS Microbiol. Lett. 307:80–86.
 41. Mitsuoka T, Hayakawa K. 1972. The fecal flora of man. I. Communication: the composition of the fecal flora of different age groups. Zentralbl. Bakteriol. Orig. A 223:333–342.
 42. Furet J-P, Firmesse O, Gourmelon M, Bridonneau C, Tap J, Mondot S, Doré J, Corthier G. 2009. Comparative assessment of human and farm animal faecal microbiota using real-time quantitative PCR. FEMS Microbiol. Ecol. 68:351–362.
 43. Nicolas-Chanoine M-H, Gruson C, Bialek-Davenet S, Bertrand X, Thomas-Jean F, Bert F, Moyat M, Meiller E, Marcon E, Danchin N, Noussair L, Moreau R, Leflon-Guibout V. 2013. 10-Fold increase (2006–11) in the rate of healthy subjects with extended-spectrum β -lactamase-producing *Escherichia coli* faecal carriage in a Parisian check-up centre. J. Antimicrob. Chemother. 68:562–568.
 44. Bettelheim KA, Faiers M, Shooter RA. 1972. Serotypes of *Escherichia coli* in normal stools. Lancet ii:1223–1224.
 45. Skurnik D, Bonnet D, Bernède-Bauduin C, Michel R, Guette C, Becker Balaire J-MC, Chau F, Mohler J, Jarlier V, Boutin Moreau J-PB, Guillemot D, Denamur E, Andremont A, Ruimy R. 2008. Characteristics of human intestinal *Escherichia coli* with changing environments. Environ. Microbiol. 10:2132–2137.
 46. Escobar-Páramo P, Clermont O, Blanc-Potard H, Bui A-B, Le Bougué-nec C, Denamur E. 2004. A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. Mol. Biol. Evol. 21:1085–1094.
 47. Le Gall T, Clermont O, Gouriou S, Picard B, Nassif X, Denamur E, Tenaillon O. 2007. Extraintestinal virulence is a coincidental by-product of commensalism in B2 phylogenetic group *Escherichia coli* strains. Mol. Biol. Evol. 24:2373–2384.
 48. Ostblom A, Adlerberth I, Wold AE, Nowrouzian FL. 2011. Pathogenicity island markers, virulence determinants *malX* and *usp*, and the capacity of *Escherichia coli* to persist in infants' commensal microbiotas. Appl. Environ. Microbiol. 77:2303–2308.
 49. Nowrouzian FL, Oswald E. 2012. *Escherichia coli* strains with the capacity for long-term persistence in the bowel microbiota carry the potentially genotoxic *pks* island. Microb. Pathog. 53:180–182.
 50. Adiba S, Nizak C, van Baalen M, Denamur E, Depaulis F. 2010. From grazing resistance to pathogenesis: the coincidental evolution of virulence factors. PLoS One 5:e11882. doi:10.1371/journal.pone.0011882.
 51. Diard M, Garry L, Selva M, Mosser T, Denamur E, Matic I. 2010. Pathogenicity-associated islands in extraintestinal pathogenic *Escherichia coli* are fitness elements involved in intestinal colonization. J. Bacteriol. 192:4885–4893.
 52. Mammeri H, Eb F, Berkani A, Nordmann P. 2008. Molecular characterization of AmpC-producing *Escherichia coli* clinical isolates recovered in a French hospital. J. Antimicrob. Chemother. 61:498–503.
 53. Scanvic-Hameg A, Chachaty E, Rey J, Pousson C, Ozoux ML, Brunel E, Andremont A. 2002. Impact of quinupristin/dalfopristin (RP59500) on the faecal microflora in healthy volunteers. J. Antimicrob. Chemother. 49:135–139.
 54. Andremont A. 2003. Commensal flora may play key role in spreading antibiotic resistance. ASM News 63:601–607.
 55. Grenet K, Guillemot D, Jarlier V, Moreau B, Dubourdieu S, Ruimy R, Armand-Lefevre L, Bau P, Andremont A. 2004. Antibacterial resistance, Wayampis Amerindians, French Guyana. Emerg. Infect. Dis. 10:1150–1153.

III. Commentaires de l'article n°1

Notre objectif était d'étudier la répartition des groupes phylogénétiques de *E. coli* dans la flore commensale intestinale de l'homme.

La population étudiée provient d'une étude épidémiologique prospective, le projet « Coliville ». Il s'agit de patients consultants en médecine générale. Cette étude inclut 280 sujets sains qui ont été appariés à 366 patients bactériémiques inclus dans l'étude Colibafi (Projet de recherche clinique INSERM RBM 03-58 "Physiopathologie des bactériémies à *Escherichia coli*") en termes d'âge, de sexe, et de lieu de résidence : l'Ile de France. La durée de la participation du sujet est de un jour. La durée totale de la recherche a été de 3 ans. Dix-neuf centres ont participé à cette étude. Notre étude a porté sur les 100 premiers prélèvements de l'étude Coliville. Ainsi, notre étude comprend 100 patients vus en consultation de médecine générale par des praticiens rattachés au Département de Médecine Générale de l'UFR de Médecine de l'Université Paris Diderot-Paris 7 (Pr. Michel Nougairède et Pr. Alain Eddi). Le recrutement s'est fait lors des consultations sur la base du volontariat.

Critères d'inclusion des patients :

- Patient âgé de plus de 18 ans.
- Habitant l'Ile de France.

Critères de non inclusion des patients :

- Les personnes vulnérables telles que définies dans les dispositions relatives à la Recherche Biomédicale et exposées dans le Code de la Santé Publique.
- Prise d'antibiotiques dans le mois précédent ou prescription lors de la consultation.
- Hospitalisation dans les 3 derniers mois.

- Personnes vivant en institutions, longs séjours.
- Immunodépression connue (VIH, immunosuppresseurs, chimiothérapie, corticoïdes au long cours quel que soit la dose > 2 semaines).
- Pathologies gastro-intestinales chroniques (recto-colite ulcéro-hémorragique et maladie de Crohn).

Le caractère réellement commensal de cette collection était un préalable à la qualité des résultats. Nous nous sommes efforcés d'être extrêmement rigoureux dans le recrutement des sujets afin d'écartier tout biais de sélection modifiant la flore intestinal, mais également en imposant une unité de temps : un recrutement récent de 2009 à 2012 et de lieu, la région parisienne. D'autre part, ces critères nous ont permis d'écartier des situations physiopathologiques pouvant bouleverser la flore commensale comme une hospitalisation ou une antibiothérapie antérieures, une intervention sur le tube digestif, une maladie inflammatoire chronique de l'intestin, une immunodépression ou une vie en communauté fermée.

Par ailleurs, les sujets ayant été recrutés dans 19 cabinets de médecine générale dispersés dans la région parisienne, des phénomènes de proximité épidémiologique liés à la propagation de souches ont été évités. Cette approche nous assure un échantillonnage représentatif des souches commensales. Les sujets étant variables en âge (24 à 86 ans) et de sex-ratio équilibré, présentent sans doute également une variabilité appréciable en terme de mode de vie et de régime alimentaire. Des études ultérieures prenant en compte ces paramètres pourraient être entreprises afin de tester leur influence sur le commensalisme.

Les faibles taux de résistance aux antibiotiques (tableau V) obtenus sur les 280 souches de l'étude Coliville isolés de 244 sujets dont sont extraits les 100 sujets étudiés, constitue un

argument rétrospectif de cette qualité (E. Denamur, communication personnelle). Ils sont également un élément rassurant sur le niveau de résistance acquise de la flore communautaire à *E. coli* de la région parisienne qui contraste avec la résistance détectée au cours des infections extra intestinales comme lors de l'étude Colibafi (tableau V) ou par rapport à l'extension des souches de *E. coli* porteuses de BLSE de type CTX-M.

Antibiotiques	Coliville (%)	COLIBAFI (%)	OR	<i>p</i>
AMX	29.3	65.4	4.48	< 0.001
AMC	15.7	56.4	7.71	< 0.001
FOX	1.1	5.7	7.5	0.007
CTX	1.1	5.4	7	0.01
AN	0	2.5	-	-
OFX	3.6	21.4	7.25	< 0.0001
SXT	15.4	40	3.46	< 0.0001

Tableau V : Comparaison de la résistance aux antibiotiques des souches de *E. coli* des collections Coliville (280 sujets) et COLIBAFI (366 sujets) (E. Denamur, communication personnelle). AMX : amoxicilline, AMC : amoxicilline-acide clavulanique, FOX : céfoxitine, CTX : céfotaxime, AN : anikacine, OFX : ofloxacine, SXT : cotrimoxazole.

Très peu d'études ont porté sur la quantification de *E. coli* dans les selles. La plupart sont anciennes et phénotypiques, elles retrouvent une concentration par gramme de fèces chez les humains de 10^7 à 10^9 UFC (Mitsukoa *et al.*, 1972 ; Slanetz *et al.*, 1957). Après l'âge de deux ans, cette densité se stabilise à environ 10^8 UFC par gramme jusqu'à ce qu'elle décroisse chez les plus âgés (Mitsukoa *et al.*, 1972). L'étude de Penders (2006) retrouve par une approche moléculaire de PCR en temps réel, une densité de *E. coli* de 10^9 par gramme de selles chez les nouveau-nés.

Ce travail constitue la première quantification de *E. coli* dans une population adulte saine par méthode moléculaire. Dans notre étude, le portage de *E. coli* chez l'humain est de 98 %. La prévalence selon Mitsukoa est de 90 %.

Il semble que la densité de *E. coli* baisse à l'âge adulte de 10^9 par gramme de selles chez les nouveau-nés à 10^7 chez l'adulte. Nous ne retrouvons pas de baisse de densité chez la population de plus de 60 ans.

Nos résultats montrent que le portage commensal de *E. coli* est quantitativement constant dans une population de sujets sains et valident la population étudiée (critères d'inclusion et d'exclusion) en vue de travaux ultérieurs pour l'étude du commensalisme. Nous disposons d'un véritable panel de souches commensales. Par contraste, les tests effectués sur les selles des patients hospitalisés à l'hôpital d'Avicenne (Bobigny, France) montraient une grande hétérogénéité de portage (données non publiées).

Par ailleurs, notre étude qui comportait un nombre comparable d'hommes et de femmes et des sujets d'âges variables n'a pas conduit à mettre en évidence de différence significative dans la répartition des groupes en fonction du sexe ou de l'âge des sujets. Ce résultat est contradictoire avec ceux d'autres études qui montrent une différence en fonction du sexe

(Gordon *et al.*, 2005). Mais ces études sont basées sur le typage d'un clone par individu, donc sur l'hypothèse du clone dominant.

En fait, nos résultats ne sont pas contradictoires avec cette hypothèse, puisque nous avons obtenus 78% de corrélation entre notre méthode de typage et celle d'un clone pris au hasard sur le milieu d'isolement des selles des individus. Cette corrélation est d'ailleurs illustrée sur l'analyse factorielle des correspondances (AFC) (figure 1 de l'article) qui montre la proximité entre les variables groupes phylogénétiques d'une souche prise au hasard par individu et les variables groupe phylogénétique dominant.

Cette redondance dans l'information étant critiquable sur le plan statistique ; nous avons refait l'AFC en supprimant la variable groupe phylogénétique de la souche prise au hasard par individu ainsi que les variables conservant la résistance aux antibiotiques et en ajoutant le nombre de groupes phylogénétiques détectés par individu répartis en 3 variables qualitatives : 1 groupe unique, 2 à 3 groupes et 4 groupes (Figure 12).

Le plan F1/F2 qui rend compte de 27,97% de la variance, retrouve une opposition entre 2 structures caractérisées respectivement par la prédominance ou l'exclusivité des souches du groupe B2 (valeurs négatives du premier facteur) et la prédominance des souches du groupe A (valeurs positives du premier facteur).

Cette opposition permet de retrouver la plus grande fréquence des souches de ces 2 groupes chez l'homme, les groupes B1 et D y étant minoritaires.

D'autre part, l'analyse montre la liaison entre groupe A et nombre important de groupes phylogénétiques et groupe B2 et nombre faible ou absence de groupe phylogénétiques minoritaires.

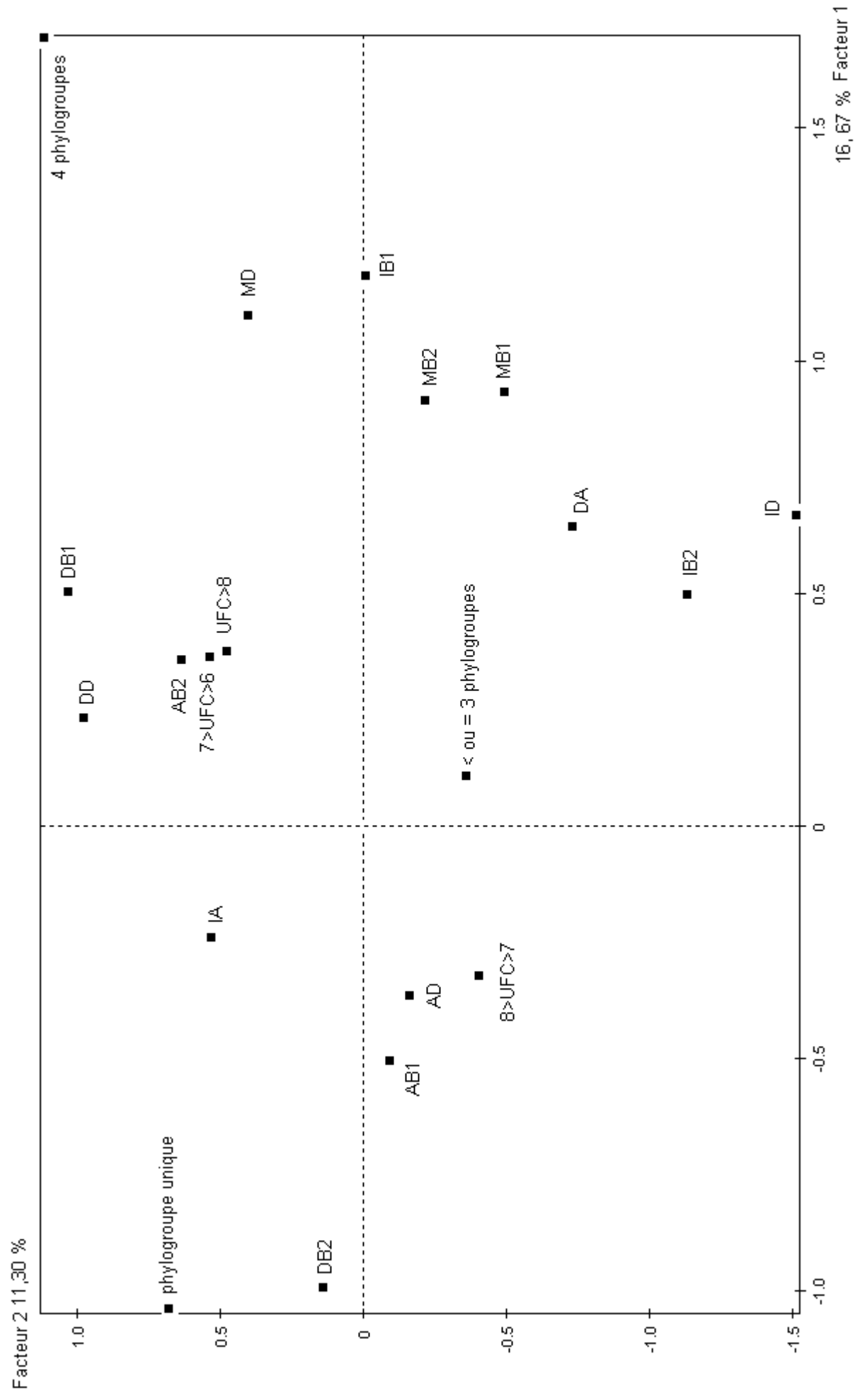


Figure 12 : Analyse factorielle des correspondances des souches de *E. coli* isolées chez les 98 sujets de l'étude. Projections sur le plan F1, F2 des groupes phylogénétiques dominants (DA, DB1, DB2 et DD), des groupes phylogénétiques intermédiaires (IA, IB1, IB2 et ID) et des groupes phylogénétiques mineurs (MA, MB1, MB2 et MD). Le nombre de groupes phylogénétiques détectés par individu (4 phylogroupes, inférieur ou égal à 3 phylogroupes, et un phylogroupe unique) ainsi que la quantité d'UFC de *E. coli* par individu (UFC > 8, 8 > UFC > 7 et 7 > UFC > 6) sont également projetés sur le plan. Les pourcentages de variance de chacun des axes sont indiqués.

Ainsi le profil d'abondance de populations B2 exclusif ou dominant est retrouvé chez 37% des individus ce qui est conforme à sa prévalence actuelle dans les pays industrialisés mais nettement différent de ce qu'il était en France et dans les régions en développement (Tableau I et II).

Cette augmentation de la prévalence des *E. coli* B2 en situation commensale n'a pas d'explication. La rareté des groupes minoritaires ou le caractère exclusif des B2 qui traduit leur caractère de « super colonisateur » ce que confirme leur plus grand équipement en facteur de colonisation et / ou de virulence pourrait expliquer cette augmentation (Nowrouzian *et al.*, 2005). Elle pourrait être une cause du succès épidémiologique des *E. coli* porteurs de BLSE CTX-M qui sont souvent des clones du groupe B2.

Les individus portant de façon exclusive des *E. coli* du groupe B2 seraient-ils plus exposés que les autres infections extra-intestinales et notamment aux pyélonéphrites ? Cette hypothèse mériterait d'être explorée dans un travail futur.

Par ailleurs, si les *E. coli* B2 sont fréquemment chez l'homme et chez les animaux sauvages, ils sont rares chez les animaux domestiques (Escobar-Paramo *et al.*, 2006). La comparaison

de leur nature en termes de sous-groupes, de types O somatiques et de facteurs de colonisation/virulence en fonction de leur origines animale ou humaine sera étudiée dans le second article.

Contrastant avec le profil d'abondance des populations B2, le profil A est caractérisé par la présence de groupes minoritaires ce qui en ferait, malgré une fréquence en terme de groupe dominant comparable à celle du groupe B2, un moins bon colonisateur de ce groupe. Il est d'ailleurs significativement moins bien équipé en facteurs de colonisation/virulence et sa fréquence a diminuée en France par rapport aux années 1980 et par rapport aux régions en développement

(Tenaillon *et al.*, 2010) au profit du B2.

Mais notre étude a montré qu'il était également plus fréquemment résistant aux antibiotiques que le groupe B2. Ce dernier trait illustre la balance résistance-virulence, mise en évidence chez les souches de *E. coli* responsables d'infections extra-intestinales (Vila *et al.*, 2002).

Nous n'avons pas clairement caractérisé par l'AFC les structures de populations B1 majoritaire et D majoritaire du fait de la rareté de ces groupes phylogénétiques chez l'homme. Ces profils pourraient être étudiés avec un échantillonnage de selles humaines plus important ou en analysant des selles d'origine animale où les B1 sont fréquents ce qui fera l'objet du second travail expérimental. Cette moindre caractérisation des profils B1 et D peut également tenir au mode de représentation statistique que nous avons utilisé avec l'AFC. Cette technique utilise des variables qualitatives discontinues avec un code binaire : 1 présent, 0 absent.

Or les pourcentages des différents groupes que nous avons déterminés sont des variables continues, quantitatives. Nous avons donc effectué une analyse en composante principale (ACP) (Figure 13) qui permet d'étudier de telles variables. Sur le plan F1, F3 de l'ACP qui représente 70.09 % de la variance, les 4 variables sont projetées selon 4 vecteurs bien distincts

ce qui illustre le fait que les groupes B1 et D sont également retrouvés majoritaires chez l'homme (dans 13 et 11% des cas respectivement) (Figure 12).

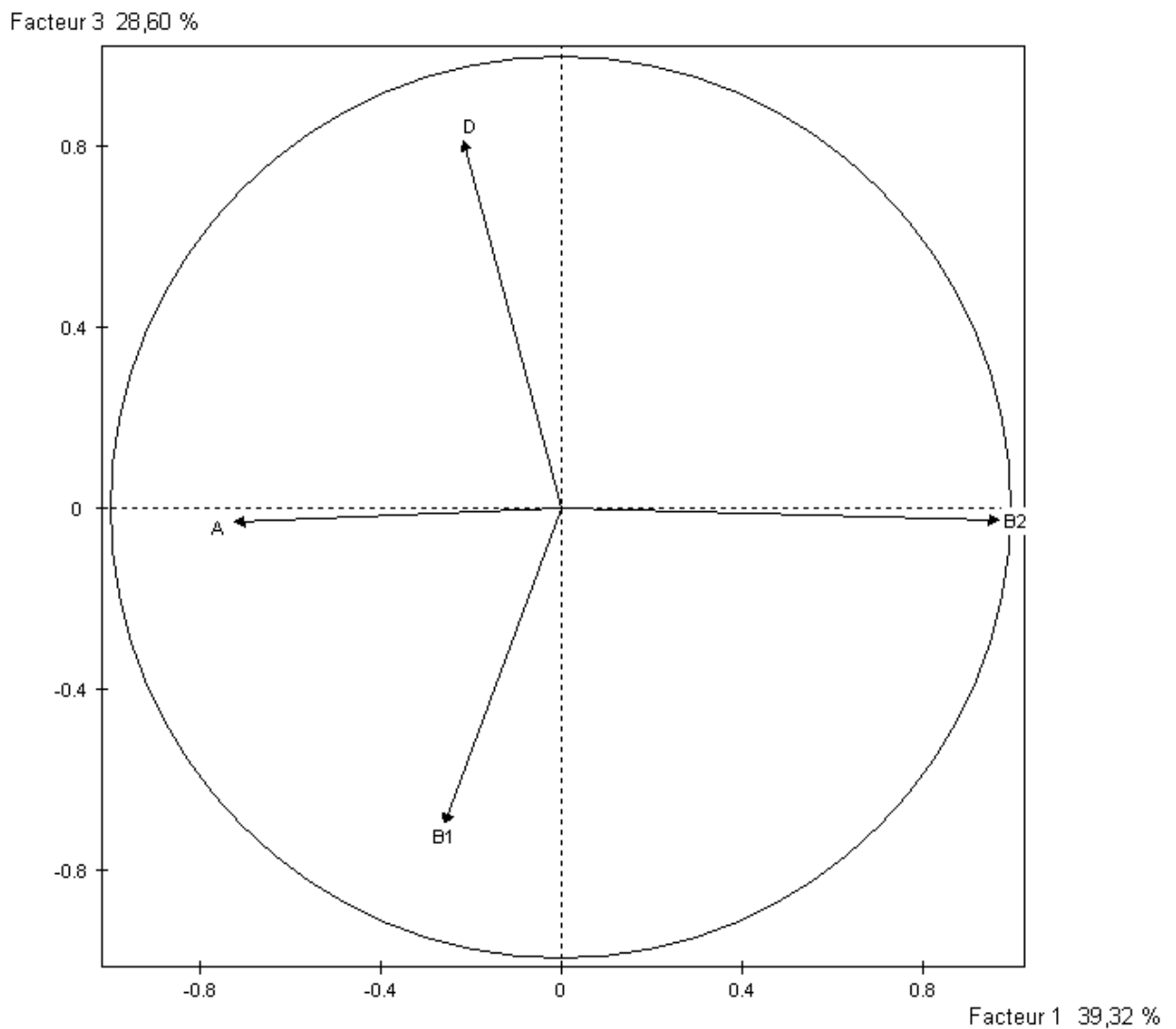


Figure 13 : Analyse en composantes principales des pourcentages des 4 groupes phylogénétiques des souches de *E. coli* d'origine humaine. Projections des 4 vecteurs correspondant aux groupes phylogénétiques A, B1, B2 et D sur le plan F1, F3 de l'analyse. Les pourcentages de variance de chacun des axes sont indiqués.

Si l'on considère la projection des individus sur ce plan F1/F3 (Figure 14), on voit qu'ils se répartissent selon les 4 profils d'abondance des populations commensales qui se projettent sur les valeurs fortes positives et négatives de l'axe F1 et de l'axe F3.

Les valeurs très positives de F1 distinguent les individus présentant le profil B2 exclusif ou majoritaire qui s'opposent aux individus à profil A majoritaire (valeurs négatives de F1). Tout au long du facteur 1, les individus se répartissent selon un gradient de profils variables. Cette diversité importante sur une série limitée de 98 sujets illustre le polymorphisme du portage commensal à *E. coli*. Ainsi, à côté du clone dominant, les hommes hébergent des populations de *E. coli* minoritaires qui peuvent expliquer l'adaptation de l'espèce à des bouleversements physiologiques ou physiopathologiques ainsi que l'émergence de clones résistants. Cela montre également la valeur de la technique quantitative de typage que nous avons développé qui permet de détecter des groupes en proportions minoritaires et qui est applicables à des séries importantes d'individus.

Cependant, si *E. coli* est bien adapté à l'homme vivant en région parisienne, omnivore, médicalisé et ayant une alimentation industrielle, comment cet équilibre entre les groupes phylogénétiques se présente et se maintient quand on change d'espèce d'hôte, de régime et de mode de vie. C'est l'objet du second travail expérimental.

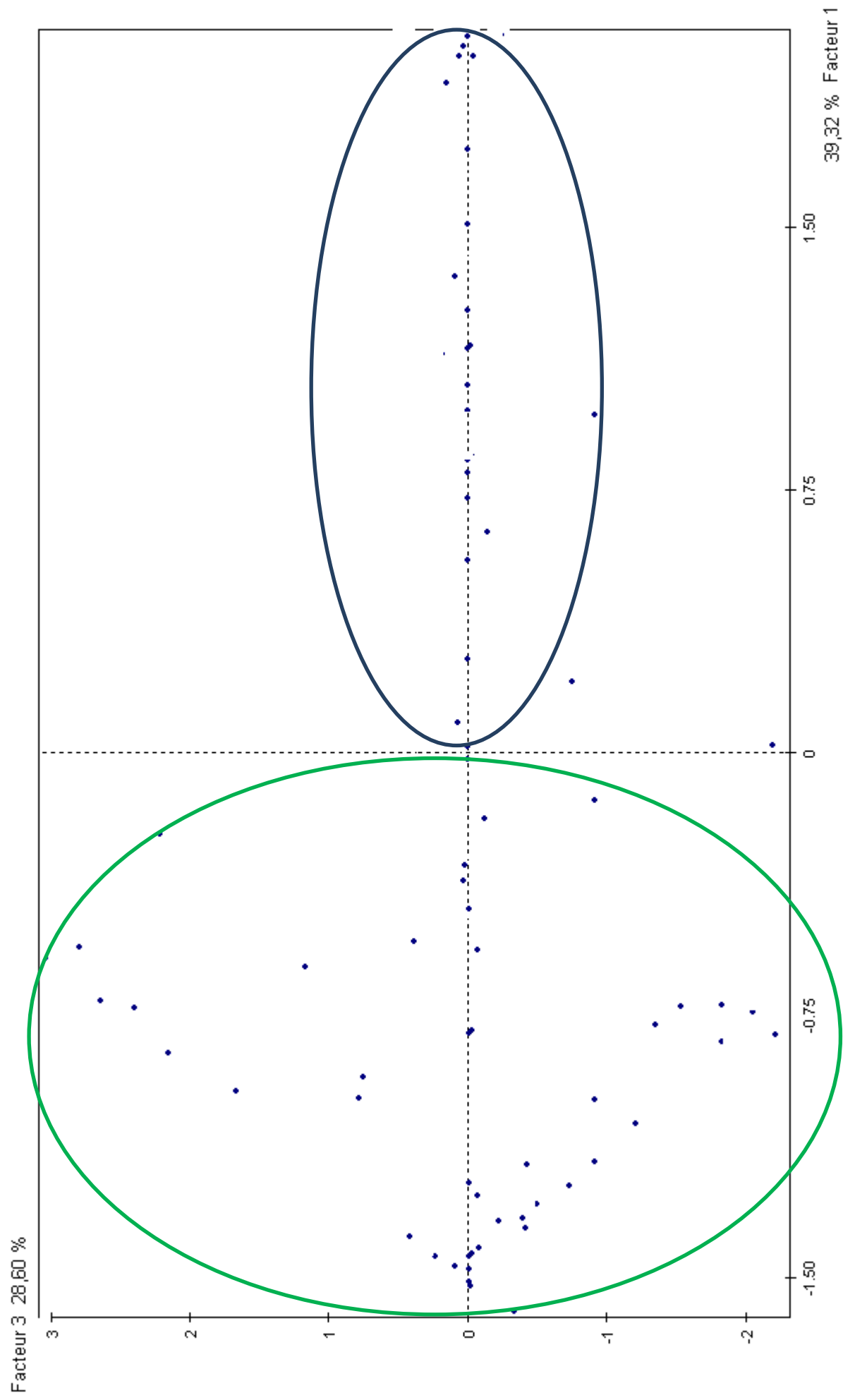


Figure 14 : Analyse en composantes principales des pourcentages des 4 groupes phylogénétiques des souches de *E. coli* d'origine humaine. Projections des 98 individus sur le plan F1, F3 de l'analyse. Cercle bleu : profil B2 majoritaire, cercle vert : profil A majoritaire.

ARTICLE N°2

I. Introduction de l'article n°2

Nous avons vu que *E. coli* était effectivement adapté à l'espèce humaine; cette adaptation s'exprimant par la quasi constance de sa présence (98% dans notre série) et par les taux élevés pour une bactérie aéro-anérobie de ce portage (Smati *et al.*, 2013).

D'autre part, nous avons mis en évidence une dualité entre un profil d'abondance des populations commensales caractérisé par la prédominance du groupe phylogénétique B2 qui apparaît comme un super-colonisateur et un profil caractérisé par la prédominance du groupe phylogénétique A qui laisse des proportions appréciables des autres groupes phylogénétiques présents et apparaît plus conforme à une notion d'équilibre intra-spécifique.

Par ailleurs, la plus grande sensibilité de la technique de qPCR par rapport aux techniques de détection phénotypique nous a permis de mettre en évidence des groupes ultra minoritaires chez plus de 40% des individus qui peuvent être potentiellement porteurs de tel ou tel trait (virulence, résistance aux antibiotiques) et émerger à l'occasion d'une pression de sélection.

Cependant, *E. coli* est également un commensal des animaux à sang chaud où sa présence à déjà été mise en évidence lors de plusieurs études (Gordon et Cawling 2003, Escobar-Paramo *et al.*, 2006).

Aborder la caractérisation des proportions des différents groupes phylogénétiques des souches commensales d'espèces animales nous est apparu répondre à la problématique de l'adaptation.

En effet, il s'agit de tester l'influence de paramètres établis : longueur du tube digestif, régime alimentaire, mode de vie, espèce (qui conditionne le terrain immunogénétique) sur l'équilibre quantitatif des populations de *E. coli*.

Par ailleurs, l'étude effectuée chez l'homme ne nous a permis de mettre en évidence que 2 structures caractérisées par la prédominance des groupes B2 ou A ce qui est conforme à la plus grande fréquence de ces groupes chez l'homme.

Si les souches du groupe B2 sont également fréquentes chez les animaux sauvages, elles sont rares chez les animaux domestiques, pourtant épidémiologiquement plus proches de l'homme.

Cet apparent paradoxe justifie la comparaison des structures de populations de type B2 d'origine humaine et d'animaux sauvages. Les souches B2 de ces derniers sont-elles, comme chez l'homme, des « super-colonisatrices » (Diard *et al.*, 2010, Nowrouzian *et al.*, 2005) et pour ce faire sont-elles suréquipées en facteurs d'établissement et potentiellement de virulence ?

L'analyse des *E. coli* commensaux peut également nous permettre de tester l'existence de structures B1 ou D dominant et de connaître leur répartition dans les espèces.

Un autre objectif est de tester l'hypothèse du clone dominant établi chez l'homme et décrit comme s'installant dès les premiers jours de la vie (Nowrouzian *et al.*, 2005). En revanche, cette hypothèse est à vérifier dans les différentes espèces animales.

Enfin, si des profils d'abondance spécifiques d'origine animale pouvaient être mis en évidence, ils pourraient servir de marqueurs pour tracer l'origine animale ou humaine d'une contamination fécale de l'environnement et notamment de l'eau (Raith *et al.*, 2013, Carlos *et al.*, 2010).

Nous avons donc étudié les proportions des 4 groupes phylogénétiques principaux de *E. coli* dans les selles de différentes espèces d'animaux domestiques et sauvages vivant dans la région parisienne en utilisant la technique de PCR en temps réel mise au point pour les selles d'origine humaine.

Quantitative analysis of commensal *Escherichia coli* populations reveals host specific enterotypes at the intra-species level

Mounira Smati^{1,2,3}, Olivier Clermont^{1,4}, Frédéric Fourreau³, Anthony David¹, Anne-Sophie Daubié^{1,3}, Cécile Hignard⁵, Odile Loison⁵, Bertrand Picard^{1,2,3} and Erick Denamur^{1,4}.

¹INSERM, IAME, UMR 1137, F-75018 Paris, France

²Univ Paris Nord, IAME, UMR 1137, Sorbonne Paris Cité, Bobigny, France

³APHP, Hôpitaux Universitaires Paris Seine Saint-Denis, Hôpital Avicenne, Service de Bactériologie-Virologie-Hygiène, Bobigny, France

⁴Univ Paris Diderot, IAME, UMR 1137, Sorbonne Paris Cité, F-75018 Paris, France

⁵Univ Paris Diderot, Station d'Ecologie Forestière, Sorbonne Paris Cité, Fontainebleau, France

Manuscript préparé pour envoi à Environmental Microbiology.

ABSTRACT:

The primary habitat of the *Escherichia coli* species is the gut of numerous species of warm blooded vertebrates. The species, mainly clonal, is structured into four main phylogenetic groups A, B1, B2 and D. We estimated the relative proportions of these phylogroups in the feces of 137 animals with various diets and habitats in the region of Ile-de-France by a real-time PCR method, previously developed for human. As described for enterotypes at the gut microbiota inter-phyla level, diet and habitat have a deep impact on *E. coli* enterocolitypes (i.e distinct combinations of relative proportions of phylogenetic groups). Three main animal enterocolitypes were distinguished, characterized by the B1, B2 and A dominant phylogroup, respectively and by distinct diet or host habitat. Moreover, the phylogroup B2 was never exclusive in the animal panel, unlike humans. This leads us to analyze further 46 human and animal B2 strains according to their exclusive or not character in feces. The B2 strains were different between human and animal strains according to their O-types and B2 subgroup belonging, however two genes *sfa/foc* and *clbQ*, showed a strong association with the exclusive character of the strain only in humans. These results argue for a major capacity of colonization which is specific to human B2 strains. This conducts to distinguish a fourth and particular human enterocolitype.

INTRODUCTION

Gut microbiota in vertebrates often outnumber host cells, and is required for development and homeostasis in adult life. It has critical roles such as fermentation of non-digestible dietary components in the large intestine, and maintaining balance with the host's metabolism and immune system (Flint, 2012). The bacterial communities codiversified with their hosts and their composition depends of numerous parameters including host phylogeny, gut physiology and host diet (Ley, 2008). The gut microbiota of humans living a modern life-style is typical of omnivorous primates. Recently, three robust clusters called enterotypes, corresponding to characteristic community structures, were identified in the human gut microbiome (Arumugam *et al.*, 2011). These enterotypes are mostly driven by species composition at the genus level and represent clusters of relative abundance profiles. These three enterotypes were also found in chimpanzees, supporting the hypothesis that enterotypic variation was present in great apes before the human-chimpanzee divergence (Moeller *et al.*, 2012). Whereas strict anaerobes constitute the large majority of this microbiota, *Escherichia coli* is the predominant aerobic organism in the gastrointestinal tract. It is an inhabitant of the intestines and faeces of warm-blooded animals and reptiles (Gordon and Cawling 2003). In human, it is found around 10^8 CFU per gram of faeces. This implies that in microbiota and microbiome analyses using 16S rRNA and whole-genome shotgun sequence data, respectively, the *Escherichia* genus is at the threshold of detection (Claesson *et al.*, 2012; Arumugam *et al.*, 2011).

The highest prevalence of *E. coli* is found in humans (more than 90%) (Smati *et al.*, 2013, Lescat *et al.*, 2013) followed by domestic (64%) and wild (45%) animals (Lescat *et al.*, 2013). Also, *E. coli* is more frequent in mammals (56%) than in birds (23%) and reptiles (10%) (Tenailon *et al.* 2010, Gordon and Cawling 2003). Further analyses of the intra-specific genetic structure of the *E. coli* species have shown that it can be considered as a clonal species (Desjardins *et al.*, 1995) with the delineation of at least four main phylogenetic groups designated A, B1, B2 and D (Ochman and Selander 1984, Herzer *et al.*, 1990). Interestingly, these groups differ in their commensal ecological niches. Commensal human strains in Western countries belong mainly to the B2 phylogenetic group (43%), followed by A (24%), D (21%) and B1 (12%) phylogroups. Domesticated animals have a decreased proportion of B2 strains than their wild counterparts and an increased proportion of A strains (Escobar-Paramo *et al.*, 2006, Baldy-Chudzik *et al.*, 2008). It seems that the factors that shape the

global microbiota as gut anatomy and physiology and diet are also involved in the prevalence of *E. coli* and its various phylogenetic groups (Gordon Cawling 2003, Tenaillon *et al.*, 2010). However, most of the works on *E. coli* commensal strain epidemiology are based on the study of one (considered as the dominant) or few clones. Recently, using a real-time PCR method quantifying both the total numbers of *E. coli* and of the 4 main phylogenetic groups, we showed a high within-individual diversity of the various phylogenetic groups in human healthy subjects. Two distinct population structures of relative abundance of phylogroups were evidenced: the dominant phylogenetic group B2, which excluded the coexistence in faeces of other phylogroups and the dominant phylogroup A, which allows more diversity (Smati *et al.*, 2013). The B2 phylogroup is of particular interest as strains from this phylogroup are increasingly found in human as commensal in Western countries (Escobar-Paramo *et al.*, 2004, Duriez *et al.*, 2001, Tenaillon *et al.*, 2010) and are involved in extra-intestinal infections (Picard *et al.*, 1999). This virulence is linked to numerous factors such as adhesins, iron capture systems, toxins and protectins (Johnson and Stell 2000). As these factors have also been reported in commensal strains (Nowrouzian *et al.*, 2005), it has been proposed that extra-intestinal virulence could be a by-product of commensalism (Diard *et al.*, 2010, Adiba *et al.*, 2010). Furthermore, *E. coli* strains of phylogenetic group B2 harbour a genomic island called "*pks*" that codes for the production of a polyketide-peptide genotoxin, colibactin. Colon colonization with *E. coli* strains harbouring the *pks* island could contribute to the development of sporadic colorectal cancer (Cuevas-Ramos *et al.*, 2010).

As very little is known about quantitative data of the commensal *E. coli* microbiota in animals, we have determined by real-time PCR approach the total amount of *E. coli* as well as the proportions of the major *E. coli* phylogenetic groups in the faeces of various healthy animal species chosen for their distinct habitat, diet and proximity to humans. Furthermore, B2 phylogenetic strains of a subset of animals were more thoroughly characterised and compared to their human counterparts.

MATERIALS AND METHODS

***E. coli* isolation and identification.**

We performed sampling campaigns in October 2011-February 2012 in the Fontainebleau Forest, Ile de France, France, and in different farms of the same region. Depending on the host species being sampled, samples were collected from freshly deposited faeces (domestic animals and wild rabbits) or directly from faecal material in the rectum of the hunted hosts just after death (deer and boars during species regulation organized by the “Office National des Forêts” in Fontainebleau forest). Except for wild rabbits, physical examination concluded to a healthy status of the animals. Ninety domestic animals (15 pigs, 15 horses, 15 sheep, 15 cows, 15 rabbits and 15 chickens) and 47 wild animals (15 wild rabbits, 15 boars and 17 deer) were sampled by this way. The chicken was considered as omnivorous. Faecal collection was carried out using a swab/transport tube system containing Aimes transport medium. The swabs were discharged in glycerol stock solution (Cryobank; Biovalley, Marne La Vallée, France) and stored at -80°C until they were used for DNA extraction. The stool-containing suspensions were plated on chromogenic plates (Uriselect®; BioRad). Pink colonies (one per subject) were randomly picked, tested for their indole production, identified as *E. coli* if positive (according to the manufacturer’s recommendations) and stored in glycerol stock solution. Fifteen humans from the Coliville study (Smati *et al.*, 2013) were randomly selected using a table of random numbers (Kendall and Babington Smith, 1939) and studied in parallel. In addition, in cases of A dominant phylogroup by qPCR, stools were plated onto Drigalski agar plates and lactose negative colonies were sampled. *E. fergusonii* strains were identified by API20E gallery (bioMérieux, Marcy l'Etoile, France).

Quantitative-PCR assay.

DNA extraction from feces.

The QIAamp DNA Stool Minikit (Qiagen, Courtabœuf, France) was used to extract DNA from 200 mg of frozen stool sample according to the manufacturer’s recommendations with modifications (Smati *et al.*, 2013). The DNA was eluted in a final volume of 200 μl and stored at -80°C .

Quantitative PCR assay.

We used a real-time PCR assay, as previously described, targeting the genes involved in the classical phylogrouping triplex PCR method (Clermont *et al.*, 2000), to directly quantify the four main *E. coli* phylogroups (A, B1, B2 and D) from faeces. In this assay, the phylogroups

B1 (*chuA* negative, *yjaA* negative and TspE4.C2 positive), B2 (*chuA* positive, *yjaA* positive and TspE4.C2 variable) and D (*chuA* positive, *yjaA* negative and TspE4.C2 variable) were quantified by real-time PCR using probes specifically targeting B1 TspE4.C2, B2 *chuA* and D *chuA* sequences, respectively (Smati *et al.*, 2013). The proportion of each phylogroup was calculated as a percentage of the total *E. coli* count obtained by quantitative PCR for 16S rDNA (Penders *et al.*, 2006). The proportion of phylogroup A, which encompasses A₀ subgroup (*chuA*, *yjaA* and TspE4.C2 negative) and A₁ subgroup (*chuA* negative, *yjaA* positive and TspE4.C2 negative) was calculated by subtracting the proportions of phylogroups B1, B2 and D from the total *E. coli* count. The proportion of A₁ subgroup was confirmed using an *yjaA* specific probe (pyjaAA1/B2), which quantifies the A₁ subgroup and the B2 phylogroup using the following formula: $A_1 = yjaA_{\text{positive}} - B2$. The threshold of phylogroup A detection by subtraction is 15% of the total *E. coli* count. The threshold of detection of B1, B2 and D phylogroups, due to specific probes, is 0.1% of the *E. coli* population (Smati *et al.*, 2013).

Genotypic characterisation of *E. coli* colonies and colicin production.

The phylogenetic group belonging of the randomly picked *E. coli* strains were determined by the quadruplex PCR method (Clermont *et al.*, 2013) which allows detection of seven phylogroups (A, B1, B2, C, D, E, F) belonging to *E. coli sensu stricto*, of *Escherichia albertii* and of *Escherichia* clades. Furthermore, assignation of *Escherichia* clades was done by PCR as in (Clermont *et al.*, 2011).

The *E. coli* B2 strains were further sub-typed at the sequence type complex (STc) level by allele-specific PCRs as previously described (Clermont *et al.*, 2014), and O-typed by PCRs as in (Clermont *et al.*, 2007). The presence of 20 virulence factors representatives of the main classes of identified *E. coli* extraintestinal virulence determinants including adhesins (*papC*, *papG*, including *papG* alleles, *sfa/foc*, *iha*, *hra* and *ibeA*), iron capture systems (*fyuA*, *irp2*, *iroN*, *iucC* and *ireA*), protectins (*kpsE*, *neuC*, chromosomal *ompT* and *traT*) and toxins (*hlyC*, *cnf1*, *usp*, *sat* and *clbQ*) (Lefort *et al.*, 2011, Clermont *et al.*, 2011b, Johnson *et al.*, 2008). For each isolate, a virulence score, defined as the number of virulence factors present over the 20 tested, was calculated (adapted from Lefort *et al.*, 2011).

The production of colicins and phages by B2 strains was detected by plaque lysis assays, with *E. coli* K-12 (phylogroup A) used as the sensitive strain (Riley and Gordon, 1992, Gordon and O'Brien, 2006). Briefly, 10 µl of an overnight (O/N) culture in LB medium of each strain was spotted onto LB agar medium containing mitomycin (25 µM) onto which an O/N culture of

the *E. coli* K-12 strain diluted to 0.5 McFarland unit had already been plated. After O/N culture at 37°C, the result of the assay was considered to be positive for strains surrounded by a halo, corresponding to inhibition of the growth of the sensitive strain (adapted from Bleibtreu *et al.*, 2013).

Statistical analyses.

Factorial analysis of correspondence (FAC) was used to describe associations among the data; two-ways tables being analyzed using SPAD.N software (Cisia, Saint-Mandé, France). Two tables were used. The first table had 123 rows corresponding to the 123 animals with *E. coli* identified in their stool samples and 33 columns (corresponding to the 32 variables). The 32 variables were, for each group A, B1, B2 and D, the dominant phylogenetic groups, the intermediate phylogenetic groups, the minor phylogenetic groups and the absence of the phylogenetic groups (see the definitions below) as determined by qPCR; the presence of a high (more than 10^7 CFU per gram), intermediate (between 10^6 and 10^7 CFU per gram) and low (less than 10^6 CFU per gram) quantitative level of *E. coli* in faeces, the high genetic diversity (corresponding to 4 distinct phylogenetic groups per individual) and the low genetic diversity (corresponding to 3 or less distinct phylogenetic groups per individual), the animal species (horse, cow, sheep, pig, wild rabbit, domestic rabbit, deer, boar, chicken); the animal habitat (wild and domestic), and the diet (herbivorous, omnivorous).

The second table had 46 rows corresponding to 46 phylogenetic group B2 strains (31 strains of human origin and 15 strains of animal origin) and 31 columns (corresponding to the 31 variables). The 31 variables were human origin and exclusive group B2, human origin and non exclusive group B2, animal origin, the 5 B2 subgroups, the non-groupable B2 subgroup, the 20 VF genes (see table 3), the virulence score of 9 or more.

We used a binary code for each variable: present, 1; absent, 0. The FAC uses a covariance matrix based on χ^2 distances. The computation determines a plane defined by two principal axes of the analysis. The first axis, F1, accounts for most of the variance, and the second axis, F2 (orthogonal to F1), accounts for the largest part of the variance not accounted for by F1 (Greenacre 1992). Significance was assessed using a χ^2 test with a threshold <0.05 .

RESULTS

Characteristics of the studied populations.

A faeces animal collection of Ile de France was constituted during this study. In total, 137 animal fresh faeces were collected, either from deposited material on the ground or collected directly in the rectum. Among these animals, 90 were domestic (15 pigs, 15 horses, 15 sheep, 15 cows, 15 rabbits and 15 chicken) whereas 47 were wild (15 wild rabbits, 15 boars and 17 deer). By distinguishing animals on their diet, we have sampled 92 herbivorous animals (sheep, horses, cows, rabbits and deer) and 30 omnivorous animals (boars and pigs), excluding chicken.

Prevalence and quantification of *E. coli* in stools of domestic and wild animals compared to human.

First, a global quantification of *E. coli* in faeces of the 137 studied animals was done using qPCR as previously described (Penders *et al.*, 2006). The prevalence of *E. coli*, defined as the fraction of hosts in which a particular species is present, among host species was of 100% in all species except domestic rabbits (66%) (Table 1).

Differences in *E. coli* quantification were found among species (Table 1). The *E. coli* values were the lowest in wild rabbits (5.49 +/- 1.45 log CFU of *E. coli* per gram of faeces), whereas the highest value was found in humans (7.89 +/- 0.52 log CFU of *E. coli* per gram of faeces). By analyzing data according to the diet, there is a very strong impact of this factor on the total count of *E. coli* (omnivorous 7.29 *versus* herbivorous 6.13, $p < 0.001$). This effect is present even by excluding humans (omnivorous animals 6.99 *versus* herbivorous animals 6.30, $p = 0.0002$). The effect of domestication, even being significant ($p = 0.01$), is weaker (domestic animals 6.76 *versus* wild animals 6.23).

Validation of the qPCR method for animal faeces.

We determined, for the 132 animal stools where *E. coli* was detected, the respective proportions of the four main phylogenetic groups A, B1, B2 and D by the recently reported qPCR assay (Smati *et al.*, 2013). We have shown in our previous study on healthy human faeces that in 78% of cases, the dominant phylogenetic group determined by qPCR is identical to the one determined by classical triplex PCR (Clermont *et al.*, 2000) on one randomly selected clone, the discrepancies being mainly explained by randomisation. However, *E. coli* diversity in animal faeces is different from human one, with more strains

belonging to E phylogroups (Clermont *et al.*, 2011) or *Escherichia* clades (Clermont 2011 *et al.*, Walk *et al.*, 2009) for example. Also, *E. fergusonii* and *E. albertii* have been reported in animals (Herraez *et al.*, 2005, Hariharan *et al.*, 2007, Oh *et al.*, 2011), our qPCR assay identifying these strains as belonging to A phylogroup (Smati *et al.*, 2013).

To validate our assay on animal stools, one randomly selected clone per animal, considered as the dominant clone, was studied by the quadriplex PCR method, which discriminates the strains in seven main *E. coli* phylogenetic groups (A, B1, B2, C, D, E and F), and allows to distinguish *E. albertii* (Clermont *et al.*, 2013), completed by the *Escherichia* clade detection PCR (Clermont *et al.*, 2011) (data not shown). Three strains, phenotypically indistinguishable from *E. coli*, were detected in boars and deer and belong to clades III and V (Walk *et al.*, 2009, Clermont *et al.*, 2011). For the 26 samples where we found predominantly A phylogroup strains we studied also lactose-negative colonies, the *E. fergusonii* strains being identified by API20E. Only one strain of *E. fergusonii* was found in sheep. These four strains belonging to the genus *Escherichia* (one *E. fergusonii*, and three clades) were assigned as A with our qPCR method. However, this should not impact the results; as such *Escherichia* only represents a small fraction of the isolated strains (3%) (4/132). The E and F phylogroups were only found in wild animals (11 strains, 8.3% of isolates) and were assigned as D (Clermont *et al.*, 2000). The C phylogroup strains (8 strains corresponding to 6% of isolates, essentially in pigs where 50% of strains belong to this recently described phylogroup C) were assigned as A (Clermont *et al.*, 2000).

Collectively, we found 72% of concordance between the qPCR assay and the triplex PCR method on one randomly selected clone, as in human (Smati *et al.*, 2013). The rare phylogroups C, E, F and *Escherichia* clades represent less than 18% of the strains and should not impact the global results of the qPCR assay.

***E. coli* phylogroup prevalence and diversity: definition of the enterocolitypes.**

The advantage of the qPCR assay is that it can detect minor clones unlike plating (Smati *et al.*, 2013). We found a prevalence of the 4 phylogenetic groups varying from 6.7% to 100% according to the host species (Table 2). Two, three or four phylogenetic groups can be detected among animal species. Interestingly, no unique phylogroup was found. This result distinguished plainly the animals from the humans where a unique phylogroup was detected in 21% of the subjects (Smati *et al.*, 2013). The main factor affecting diversity is domestication; 55.4% of domestic animals carry the 4 main phylogroups *versus* 10.6% of wild animals ($p < 0.001$). Significant differences in prevalence of B1 and D phylogroups were

detected between domestic and wild animals (Table 2). Domestication is associated with a higher prevalence of B1 and D phylogroups (95.3% in domestic animals *versus* 29.8% in wild animals for B1 phylogroup and 74.1% *versus* 46.8% for D phylogroup, $p < 0.001$). No difference of carriage was observed for B2 and A phylogroups. There is no impact of diet on diversity in our study.

When categorizing the proportions of the phylogroups in 4 categories adapted from Schlager (Schlager *et al.*, 2002) (dominant phylogenetic group: $> 50\%$ of the population of *E. coli*; intermediate phylogenetic group: 10 to 50% of the population of *E. coli* and minor phylogenetic group $< 10\%$ of the *E. coli* population; absent phylogenetic group: undetectable), the proportions of these categories varied significantly according to the four phylogenetic groups and the hosts. This variation was complex and was described by a FAC (Figure 1).

On the F1-F2 plane, which accounted for 31.43 % of the total variance, the negative value of the first factor distinguished the dominant phylogenetic group B2, the dominant phylogenetic group D, the absence of B1 phylogenetic group, the intermediate phylogenetic group A, the intermediate phylogenetic group B2, the 2 herbivorous wild animal species (deer and wild rabbit), the wild habitat, the herbivorous diet, the low quantitative level of *E. coli* in stools (less than 10^6 CFU) and a low genetic diversity (3 or less phylogenetic groups) determining one cluster (type wild and herbivorous).

The positive values of the first factor distinguished several other variables which could be separated into 2 main groups by the positive or negative values of the second factor. Thus, the dominant phylogenetic group B1, the absence of phylogenetic group B2, the intermediate phylogenetic group D, the 4 herbivorous domestic species, horse, cow sheep and domestic rabbit, the domestic habitat, the intermediate quantitative level of *E. coli* in the stools (between 10^6 and 10^7 CFU) were distinguished by the negative values of the second factor determining a second cluster (type domestic herbivorous). Whereas the dominant phylogenetic group A, the intermediate phylogenetic group B1, the minor phylogenetic groups B1, B2 and D, the pig and chicken species, the omnivorous diet and the high quantitative level of *E. coli* in the stools (more than 10^7 CFU) and the high genetic diversity (4 phylogenetic groups) were projected on the positive values of the second factor determining a third cluster (type omnivorous). The variable boar (omnivorous and wild) was projected on an intermediate position between the first and the third cluster of the analysis.

To resume, this analysis allows us to distinguish between 3 main clusters of relative abundance of strains within the *E. coli* animal commensal populations that we propose to call

“enterocolitypes”, by analogy with the enterotypes defined in the gut microbiota (Arumugam *et al.*, 2011). The dominant phylogenetic group B2 was associated with the wild animal species, the absence of the other groups at intermediate or minor level, the low level of genetic diversity and low level of *E. coli* in faeces (enterocolitype wild and herbivorous). The dominant phylogenetic group B1 was associated with the domestic herbivorous animal species, higher level of genetic diversity and an intermediate quantitative level of *E. coli* in the stools (enterocolitype domestic herbivorous). The dominant A phylogroup is associated with variable proportions of other groups, omnivorous species and a high level of *E. coli* in stools (enterocolitype omnivorous).

As in a previous study in human feces, the exclusive character (100% of detected phylogroups) was found only in human and strongly associated to B2 phylogroup (62% of the unique phylogroups in Smati *et al.*, 2013) and appeared to be a human specific enterocolitype, this leads us to analyze more thoroughly the B2 commensal strains of human and animal origin.

Comparison of B2 commensal strains between human and animals.

We thus worked on 46 B2 phylogroup commensal strains isolated by plating either in this work on animals (15 strains) or in the preceding study on human (31 strains) (Smati *et al.*, 2013). The 15 animal isolates were found in 11 wild animals and 4 domestic animals and were never exclusive. The exclusive character of the B2 strains was found only in humans (21 exclusive strains and 10 non-exclusive ones). Consequently, we divided the B2 commensal strains into three categories: human exclusive strains, composing more than 90% of the *E. coli* microbiota, human non-exclusive strains and animal strains. The B2 strains were assigned to B2 sub-groups (sequence type complex=STc) (Clermont *et al.*, 2014) and their O-types were determined (Clermont *et al.*, 2007), both by allele-specific PCR. We also analysed their virulence factors contain and their colicin production.

The data were first described using a FAC (Fig. 2). On the plane F1-F2 of the FAC, which accounted for 25.83 % of the total variance, 3 groups of variables could be distinguished. The first group, characterized by the negative values of the 2 factors grouped the animal origin, the absence of several virulence factors, the non-groupable B2 subgroups. The second group characterized by the positive values of the second factor grouped the human and non-exclusive B2 group strains, the B2 subgroups I (STc131 according to the Achtman MLST scheme numbering), VII (STc14) and IX (STc95), the virulence factors *papGII*, *sat*, *iha*, *traT*, *ibeA*, *iucC*, *ireA*, *neuC* and *usp*. The third group was characterized by the positive values of

the first factor grouped the human and exclusive B2 group, the B2 subgroup II (STc73) and IV (STc141), the virulence factors *cnf1*, *sfa/foc*, *papGIII*, *pks*, *hlyC*, *hlyA*, *iroN* and *papC*, a virulence score of 9 or more virulence factors. The O-types were so many that they were not projected on FAC. The whole data reflect strong differences between human and animal B2 strains and at a lesser extent, between human non-exclusive and exclusive strains.

This was confirmed by the distribution of 21 virulence factors between the three populations (Table 3). The median virulence factor score was 10.47 and 9.40 respectively for exclusive and non-exclusive B2 human strains, whereas it was lower for animal isolates (6.47, $p < 0.01$). Two virulence factors were significantly lower in B2 animal strains than in B2 human strains: *papC* encoding for pilus ($p = 0.01$) and *usp* encoding for uropathogenic specific protein ($p < 0.001$). Interestingly, and despite the fact that human and animal B2 strains appears clearly distinct, two virulence factors were significantly dissimilar ($p = 0.05$) between B2 exclusive human strains and other non-dominant strains regardless their human or animal origin: *sfa/foc*, encoding for the F1c fimbriae (47.62% in human exclusive isolates versus 10% in human non-exclusive isolates and 13.37% in animal isolates) and *clbQ*, part of the *pks* island, encoding for the genotoxin colibactin (61.90% in human exclusive isolates versus 20% in human non-exclusive isolates, and 13.33% in animal isolates). The colicin production appeared similar between the three groups (data not shown).

DISCUSSION

There is a cruel lack of quantitative data on population genetic structure of *E. coli* commensal strains despite a crucial need of these data for the understanding of the normal ecology of this major facultative anaerobic bacterium of the gut.

This is mainly due to the fact that classically used quantitative methods rely on the study of plated colonies, a tedious and limited approach. We have recently developed a quantitative PCR assay that allows detection of minor clones, up to 0.1% of the total amount of *E. coli* and applied it to the human *E. coli* microbiota. In this work, we extended it to normal animal *E. coli* microbiota sampled in the same region and epoch than the human ones. Despite a greater diversity of the animal microbiota with the presence of additional *E. coli* phylogroups and associated, our method distinguishing the main *E. coli* phylogroups remains statistically relevant.

Ecological interactions between species or member species drive the relative abundances of the different phyla. It has been recently proposed that three main enterotypes could exist in humans at the total microbiota level (Arumugam *et al.*, 2011). We extended this

concept at the intra-species level by defining enterocolotypes that correspond to clusters of relative abundances of the main *E. coli* phylogroups. We identified four enterocolotypes; an exclusive B2 dominant one specific of human, and three animal ones (B2 dominant, B1 dominant and A dominant). Interestingly, the forces that drive these ecotypes are the same that was driving the total microbiota diversity, i.e. the phylogeny of the host, the diet and the domestication. Altogether, these data indicate that a complex network of interactions may exist within the intestinal microbiota from the genus to the intra-species phylogenetic groups (Trosvik *et al.*, 2014). This probably correspond to various nutrient availability coupled to specific metabolism of bacteria, cross-feeding and toxic interactions involving phages and/or colicins. Of note, we were not able to identify differences in the production of phages/colicins among the various B2 strains.

Among the *E. coli* species, B2 phylogroup strains have a particular status as they are involved in extra-intestinal infections, but also present as commensals (Nowrouzian *et al.*, 2005). It has been proposed from this data that virulence could be a by-product of commensalism (Diard *et al.*, 2010, Tourret *et al.*, 2010). Our work gains further knowledge in B2 strains as we showed that (i) B2 in wild animals and human correspond to different clones and (ii) exclusive dominant B2 strains in human are characterized by the presence of two specific VF, *sfa/foc* and *pks* (Nowrouzian and Oswald, 2012).

The gene *sfa/foc*, encoding for a subunit of F1C fimbriae did show a positive correlative association with adherence to intestinal porcine epithelial cells (Schierack *et al.*, 2013). F1C fimbriae are required for biofilm formation of *Escherichia coli* Nissle 1917, a robust colonizer of the human gastrointestinal tract. In addition, these structures are also important for persistence in infant mouse colonization (Lasaro *et al.*, 2009).

Long-term *E. coli* colonizers were significantly more likely to have the *pks* island than transient strains, which suggests that the *pks* island contributes to the pronounced gut-colonizing capacity of group B2 strains. Long-term persistence in the colon of *pks* island-containing *E. coli* strains may be associated with the induction of genomic mutations in the host intestine (Nowrouzian et Oswald 2012).

In conclusion, our work in addition to previous data (Smati *et al.*, 2013) shows that a complex network of interactions exists at several levels within the intestinal microbiota: the genus, the phylogroups within species and the clones within the species phylogroups. This leads to specific combinations of abundances levels of taxa. Interestingly, the forces that shape this structure (phylogeny of the host, diet, domestication) are acting at all levels.

REFERENCES

- Adiba, S., Nizak, C., van Baalen, M., Denamur, E., and Depaulis, F. (2010) From grazing resistance to pathogenesis: the coincidental evolution of virulence factors. *PloS One* **5**: e11882.
- Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., et al. (2011) Enterotypes of the human gut microbiome. *Nature* **473**: 174–180.
- Baldy-Chudzik, K., Mackiewicz, P., and Stosik, M. (2008) Phylogenetic background, virulence gene profiles, and genomic diversity in commensal *Escherichia coli* isolated from ten mammal species living in one zoo. *Vet Microbiol* **131**: 173–184.
- Bleibtreu, A., Gros, P.-A., Laouénan, C., Clermont, O., Le Nagard, H., Picard, B., et al. (2013) Fitness, stress resistance, and extraintestinal virulence in *Escherichia coli*. *Infect. Immun.* **81**: 2733–2742.
- Claesson, M.J., Jeffery, I.B., Conde, S., Power, S.E., O'Connor, E.M., Cusack, S., et al. (2012) Gut microbiota composition correlates with diet and health in the elderly. *Nature* **488**: 178–184.
- Clermont, O., Bonacorsi, S., and Bingen, E. (2000) Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl. Environ. Microbiol.* **66**: 4555–4558.
- Clermont, O., Christenson, J.K., Daubié, A.-S., Gordon, D.M., and Denamur, E. (2014) Development of an allele-specific PCR for *Escherichia coli* B2 sub-typing, a rapid and easy to perform substitute of multilocus sequence typing. *J. Microbiol. Methods* **101**: 24–27.
- Clermont, O., Christenson, J.K., Denamur, E., and Gordon, D.M. (2013) The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ. Microbiol. Rep.* **5**: 58–65.
- Clermont, O., Gordon, D.M., Brisse, S., Walk, S.T., and Denamur, E. (2011) Characterization of the cryptic *Escherichia* lineages: rapid identification and prevalence. *Environ. Microbiol.* **13**: 2468–2477.
- Clermont, O., Johnson, J.R., Menard, M., and Denamur, E. (2007) Determination of *Escherichia coli* O types by allele-specific polymerase chain reaction: application to the O types involved in human septicemia. *Diagn Microbiol Infect Dis* **57**: 129–136.
- Clermont, O., Olier, M., Hoede, C., Diancourt, L., Brisse, S., Keroudean, M., et al. (2011) Animal and human pathogenic *Escherichia coli* strains share common genetic backgrounds. *Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis.* **11**: 654–662.
- Cuevas-Ramos, G., Petit, C.R., Marcq, I., Boury, M., Oswald, E., and Nougayrède, J.-P. (2010) *Escherichia coli* induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 11537–11542.
- Desjardins, P., Picard, B., Kaltenbock, B., Elion, J., and Denamur, E. (1995) Sex in *Escherichia coli* does not disrupt the clonal structure of the population: evidence from random amplified polymorphic DNA and restriction-fragment-length polymorphism. *J Mol Evol* **41**: 440–448.
- Diard, M., Garry, L., Selva, M., Mosser, T., Denamur, E., and Matic, I. (2010) Pathogenicity-associated islands in extraintestinal pathogenic *Escherichia coli* are fitness elements involved in intestinal colonization. *J. Bacteriol.* **192**: 4885–4893.
- Duriez, P. (2001) Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. *Microbiology* **147**: 1671–1676.
- Escobar-Páramo, P., Grenet, K., Le Menac'h, A., Rode, L., Salgado, E., Amorin, C., et al. (2004) Large-scale population structure of human commensal *Escherichia coli* isolates. *Appl. Environ. Microbiol.* **70**: 5698–5700.

- Escobar-Páramo, P., Le Menac'h, A., Le Gall, T., Amorin, C., Gouriou, S., Picard, B., et al. (2006) Identification of forces shaping the commensal *Escherichia coli* genetic structure by comparing animal and human isolates. *Environ. Microbiol.* **8**: 1975–1984.
- Flint, H.J. (2012) The impact of nutrition on the human microbiome. *Nutr. Rev.* **70 Suppl 1**: S10–13.
- Gordon, D.M. and Cowling, A. (2003) The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology* **149**: 3575–3586.
- Gordon, D.M. and O'Brien, C.L. (2006) Bacteriocin diversity and the frequency of multiple bacteriocin production in *Escherichia coli*. *Microbiology* **152**: 3239–3244.
- Greenacre, M. (1992) Correspondence analysis in medical research. *Stat. Methods Med. Res.* **1**: 97–117.
- Hariharan, H., López, A., Conboy, G., Coles, M., and Muirhead, T. (2007) Isolation of *Escherichia fergusonii* from the feces and internal organs of a goat with diarrhea. *Can. Vet. J. Rev. Vét. Can.* **48**: 630–631.
- Herráez, P., Rodríguez, A.F., Espinosa de los Monteros, A., Acosta, A.B., Jaber, J.R., Castellano, J., and Castroa, A. (2005) Fibrino-necrotic typhlitis caused by *Escherichia fergusonii* in ostriches (*Struthio camelus*). *Avian Dis.* **49**: 167–169.
- Herzer, P.J., Inouye, S., Inouye, M., and Whittam, T.S. (1990) Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. *J. Bacteriol.* **172**: 6175–6181.
- Johnson, J.R., Johnston, B., Kuskowski, M.A., Nougayrede, J.-P., and Oswald, E. (2008) Molecular Epidemiology and Phylogenetic Distribution of the *Escherichia coli* pks Genomic Island. *J. Clin. Microbiol.* **46**: 3906–3911.
- Kendall, M.G., Babington Smith, B. (1939). Tables of Random Sampling Numbers. Cambridge, England: Cambridge University Press.
- Lasaro, M.A., Salinger, N., Zhang, J., Wang, Y., Zhong, Z., Goulian, M., and Zhu, J. (2009) F1C fimbriae play an important role in biofilm formation and intestinal colonization by the *Escherichia coli* commensal strain Nissle 1917. *Appl. Environ. Microbiol.* **75**: 246–251.
- Lefort, A., Panhard, X., Clermont, O., Woerther, P.-L., Branger, C., Mentré, F., et al. (2011) Host Factors and Portal of Entry Outweigh Bacterial Determinants To Predict the Severity of *Escherichia coli* Bacteremia. *J. Clin. Microbiol.* **49**: 777–783.
- Lescat, M., Clermont, O., Woerther, P.L., Glodt, J., Dion, S., Skurnik, D., et al. (2013) Commensal *Escherichia coli* strains in Guiana reveal a high genetic diversity with host-dependant population structure. *Environ. Microbiol. Rep.* **5**: 49–57.
- Ley, R.E. (2008) Evolution of mammals and their gut microbes. *Science* **320**: 1647–1651.
- Moeller, A.H., Degnan, P.H., Pusey, A.E., Wilson, M.L., Hahn, B.H., and Ochman, H. (2012) Chimpanzees and humans harbour compositionally similar gut enterotypes. *Nat. Commun.* **3**: 1179.
- Nowrouzian, F.L., Wold, A.E., and Adlerberth, I. (2005) *Escherichia coli* strains belonging to phylogenetic group B2 have superior capacity to persist in the intestinal microflora of infants. *J. Infect. Dis.* **191**: 1078–1083.
- Nowrouzian, F.L. and Oswald, E. (2012) *Escherichia coli* strains with the capacity for long-term persistence in the bowel microbiota carry the potentially genotoxic pks island. *Microb. Pathog.* **53**: 180–182.
- Ochman, H. and Selander, R.K. (1984) Evidence for clonal population structure in *Escherichia coli*. *Proc Natl Acad Sci USA* **81**: 198–201.
- Oh, J.-Y., Kang, M.-S., Hwang, H.-T., An, B.-K., Kwon, J.-H., and Kwon, Y.-K. (2011) Epidemiological investigation of *eaeA*-positive *Escherichia coli* and *Escherichia*

- albertii* strains isolated from healthy wild birds. *J. Microbiol. Seoul Korea* **49**: 747–752.
- Penders, J., Thijs, C., Vink, C., Stelma, F.F., Snijders, B., Kummeling, I., et al. (2006) Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* **118**: 511–521.
- Picard, B., Garcia, J.S., Gouriou, S., Duriez, P., Brahimi, N., Bingen, E., et al. (1999) The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect. Immun.* **67**: 546–553.
- Riley, M.A. and Gordon, D.M. (1992) A survey of Col plasmids in natural isolates of *Escherichia coli* and an investigation into the stability of Col-plasmid lineages. *J. Gen. Microbiol.* **138**: 1345–1352.
- Schierack, P., Rödiger, S., Kuhl, C., Hiemann, R., Roggenbuck, D., Li, G., et al. (2013) Porcine *E. coli*: virulence-associated genes, resistance genes and adhesion and probiotic activity tested by a new screening method. *PLoS One* **8**: e59242.
- Schlager, T.A., Hendley, J.O., Bell, A.L., and Whittam, T.S. (2002) Clonal diversity of *Escherichia coli* colonizing stools and urinary tracts of young girls. *Infect. Immun.* **70**: 1225–1229.
- Smati, M., Clermont, O., Le Gal, F., Schichmanoff, O., Jauréguy, F., Eddi, A., et al. (2013) Real-time PCR for quantitative analysis of human commensal *Escherichia coli* populations reveals a high frequency of subdominant phylogroups. *Appl. Environ. Microbiol.* **79**: 5005–5012.
- Tenaillon, O., Skurnik, D., Picard, B., and Denamur, E. (2010) The population genetics of commensal *Escherichia coli*. *Nat. Rev. Microbiol.* **8**: 207–217.
- Tourret, J., Diard, M., Garry, L., Matic, I., and Denamur, E. (2010) Effects of single and multiple pathogenicity island deletions on uropathogenic *Escherichia coli* strain 536 intrinsic extra-intestinal virulence. *Int. J. Med. Microbiol.* **300**: 435–439.
- Trosvik, P., de Muinck, E.J., and Stenseth, N.C. (2014) Biotic interactions and temporal dynamics of the human gastrointestinal microbiota. *ISME J.*
- Walk, S.T., Alm, E.W., Gordon, D.M., Ram, J.L., Toranzos, G.A., Tiedje, J.M., and Whittam, T.S. (2009) Cryptic lineages of the genus *Escherichia*. *Appl. Environ. Microbiol.* **75**: 6534–6544.

TABLE 1. Prevalence and quantity of *E. coli* in faeces of domestic animals, wild animals and humans.

Species	Number of stools	Prevalence by qPCR	Mean of the log10 colony forming units (CFU) value of <i>E. coli</i> per gram of feces +/- SEM	
Boar	15	100%	7.09+/-0.97	
Deer	17	100%	6.11+/-0.33	
Wild rabbit	15	100%	5.49+/-1.45	
Pig	15	100%	6.90+/-0.71	
Horse	15	100%	6.09+/- 0.77	
Sheep	15	100%	7.31+/-0.74	
Cow	15	100%	6.16+/-0.20	
Domestic rabbit	15	66%	6.82+/-0.67	
Chicken	15	100%	7.33+/-0.54	
Human (*)	15	100%	7.89+/-0.52	
Herbivorous species	92	94.5%	6.13+/-0.79	} <0.001
Omnivorous species (**)	45	100%	7.29+/-0.85	
Domestic animals	90	94.3%	6.76+/-0.80	} 0.01
Wild animals	47	100%	6.23+/-0.87	

(*) Randomly selected, from Smati *et al* 2013, (**): including humans.

TABLE 2: Prevalence of A, B1, B2 and D phylogroups among species as determined by qPCR

Species	A prevalence (%)	B1 prevalence (%)	B2 prevalence (%)	D prevalence (%)*
Boar	93.3	46.7	53.3	73.3
Deer	100	35.3	88.2	47
Wild rabbit	66.7	6.7	100	20
Pig	93.3	80	100	73.3
Horse	53.3	93.3	100	33.3
Sheep	60	100	40	100
Cow	100	100	100	80
Domestic rabbit	53.3	100	30	50
Chicken	93.3	100	66.7	100
Human	73.3	40	66.7	40
Herbivorous species	72.8	70.1	79.3	55.2
Omnivorous species**	87	53.3	73.3	62.2
Domestic animals	75.6	95.3	75.3	74.1
Wild animals	87.5	29.8	80.8	46.8

*: with a detection threshold of 15% (Smati *et al* 2013), **: including humans, ns: not significant.

TABLE 3: Prevalence of virulence factors and virulence score of human and animal *E. coli* B2 commensal strains

Genes	Description	Human exclusive (%) (n = 21)		Human non- exclusive (%) (n = 10)	Animal (%) (n = 15)	
		<i>P</i> value		<i>P</i> value		
<i>papC</i>	Pilus	42.95	0.02 ^a	40	6.67	0.01 ^b
<i>papGIII</i>	Pilus	14.28		10	6.67	
<i>sfa/foc</i>	F1C-fimbriae	47.62	0.03 ^a	10	13.37	
			0.05 ^c			
<i>iha</i>	Iron-regulated-gene-homologue adhesin	14.28		40	0	
<i>hra</i>	Heat-resistant agglutinin	23.80		10	40	
<i>ibeA</i>	Invasion of brain endothelium	42.95		40	40	
<i>fyuA</i>	Yersiniabactin receptor	95.24		90	73.33	
<i>irp2</i>	Yersiniabactin synthesis	95.24		90	73.33	
<i>iroN</i>	Salmocheilin receptor	80.95		60	46.67	
<i>aer</i>	Aerobactin synthesis	42.85		20	20	
<i>ireA</i>	Iron responsive element	23.80		40	33.33	
<i>kpsMTII</i>	Group II capsule antigen	61.90		70	60	
<i>neuC</i>	K1 capsular polysaccharide	61.90	0.04 ^a	50	26.67	
<i>ompT</i>	Outer membrane protein	80.95		90	86.67	
<i>traT</i>	Transfert protein	52.38		80	53.33	
<i>hlyC</i>	Haemolysin	33.33		20	6.67	
<i>cnf1</i>	Cytotoxic necrotising factor 1	28.57	0.03 ^a	10	0	
<i>usp</i>	Uropathogenic specific protein	95.24	0.004 ^a	100	40	<0.001 ^b
<i>sat</i>	Secreted autotransporter toxin	14.28		30	0	
<i>clbQ</i>	Genotoxin colibactin	61.90	0.006 ^a	20	13.33	
			0.05 ^c			
VF score*	As in Lefort <i>et al</i> 2010, calculated on 20 VFs	10.47+/-2.46		9.40+/-2.67	6.47+/-2.92	<0.001 ^b

^a: p significant between human exclusive and animal isolates

^b: p significant between human and animal isolates

^c: p significant between human exclusive and human non exclusive isolates

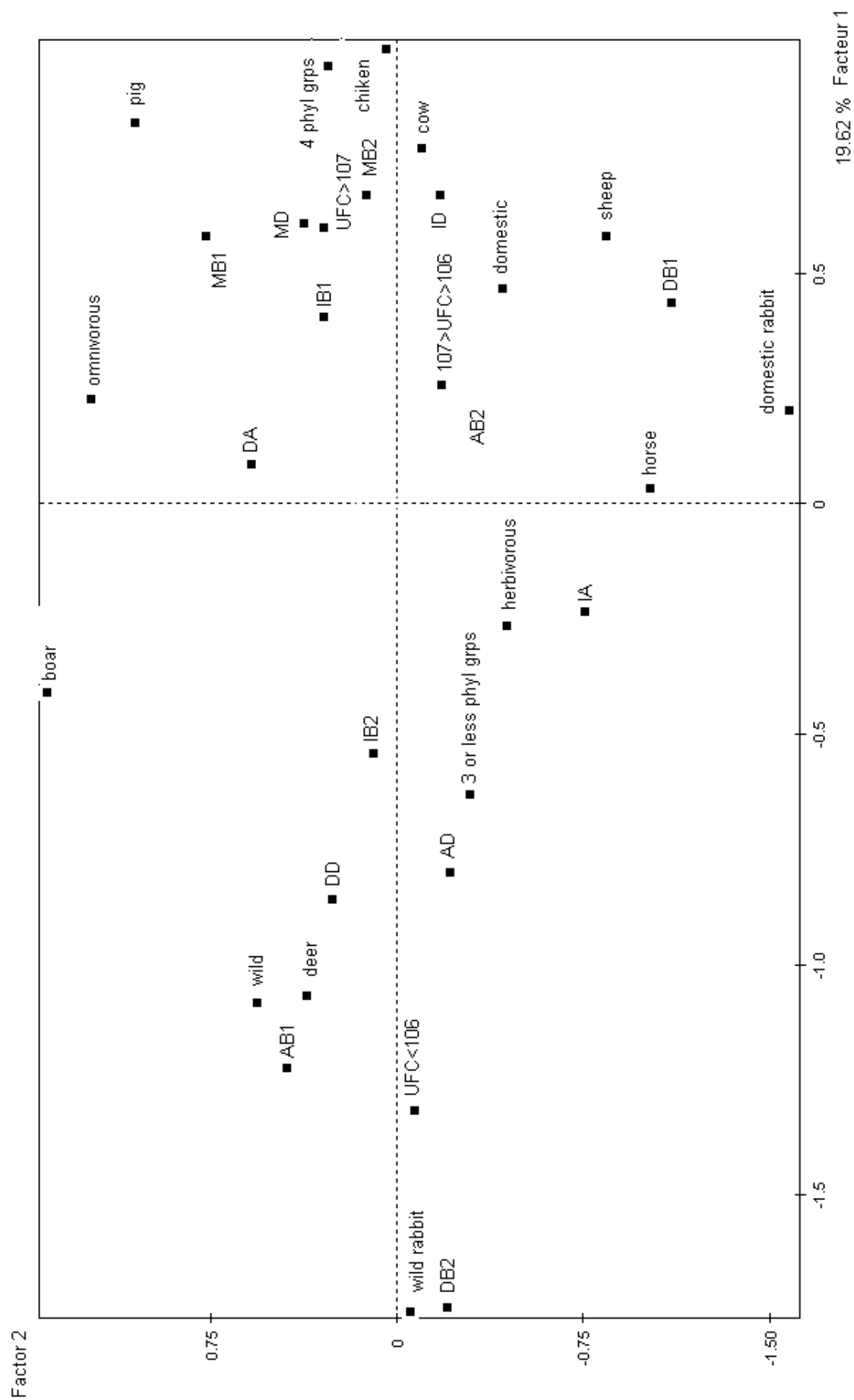


FIGURE LEGEND

FIGURE 1. Factorial analysis of correspondence (FAC) for the 123 animals with *E. coli* found in their feces. Projections on the plane F1-F2 of the dominant phylogenetic groups (DA, DB1, DB2 and DD), of the intermediate phylogenetic groups (IA, IB1, IB2 and ID), of the minor phylogenetic groups (MB1, MB2 and MD), of the absence of the phylogenetic groups (AB1, AB2 and AD), of the high genetic diversity (4 phyl grps), of the lower genetic diversity (3 or less phyl grps), of the abundance of *E. coli* in stools ($CFU > 10^7$, $10^7 < CFU < 10^6$ and $CFU < 10^6$), of the animal species (horse, cow, sheep, wild rabbit, domestic rabbit, deer, pig, boar and chicken), of the habitat (wild and domestic), of the diet (herbivorous and omnivorous).

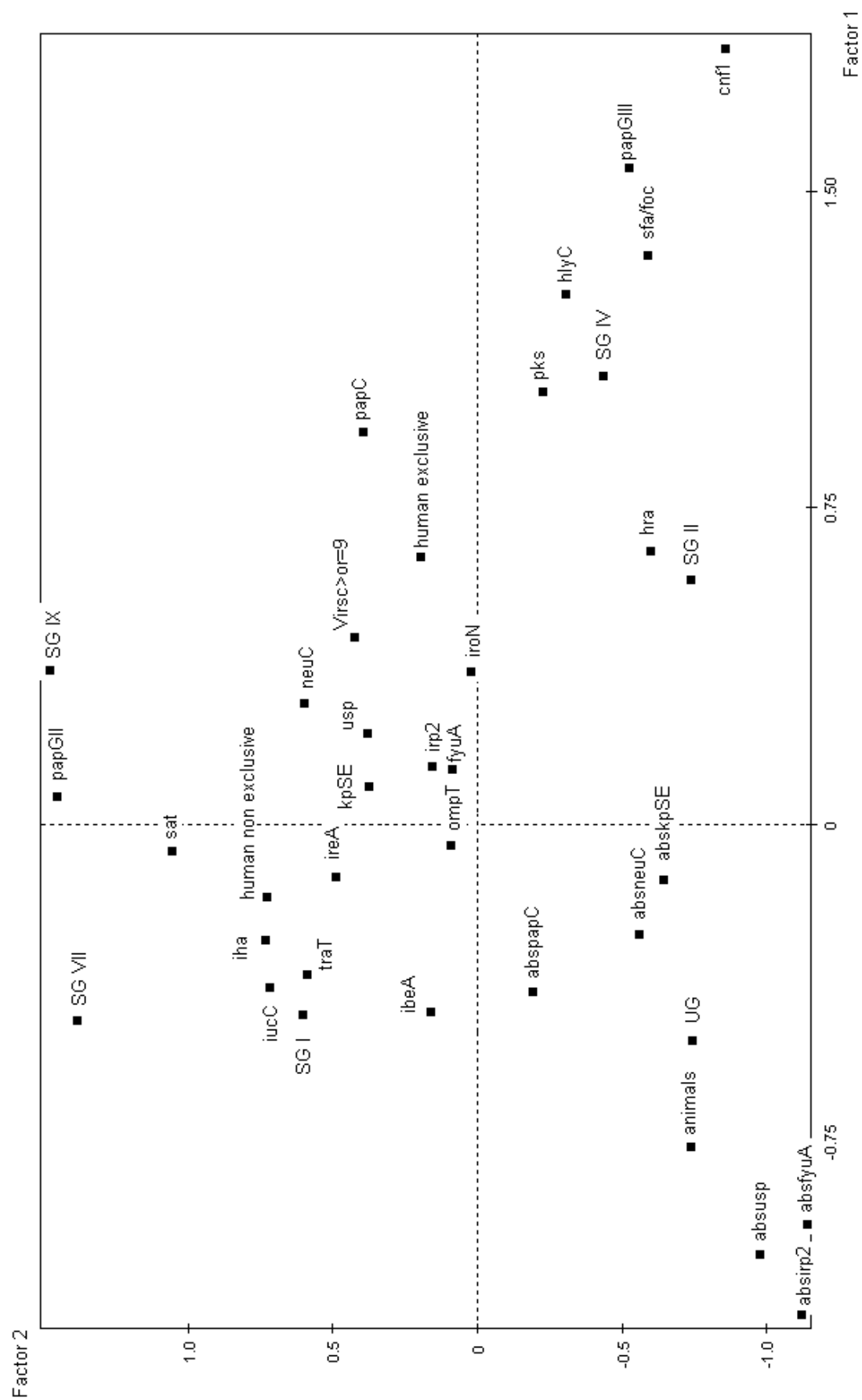


FIGURE 2.

FAC for the 46 phylogroup B2 commensal strains of human and animal origins. Projections on the plane F1-F2 of the variables: human and exclusive B2 group (humexcl), human and non exclusive B2 group (humnoex), animal origin (animals), of the B2 subgroups (SG) I, II, IV, VII and IX, of the non groupable B2 subgroups (UG), of the VFs described in table 3, of the absence (abs) of several VFs, of the virulence score of 9 or more ($\text{Virsc} \geq 9$).

III. Commentaires de l'article n°2

La technique de détection de la présence de *E. coli* dans les selles des différentes espèces animales par qPCR a présenté une plus grande sensibilité que les techniques conventionnelles par isolement.

Ainsi, nous avons mis en évidence *E. coli* chez 100% des animaux à l'exception des lapins domestiques ou cette détection a été de 66%. Ce résultat est conforme à celui obtenu par Furet *et coll.* (Furet *et al.*, 2009) dans des espèces comparables mais il est supérieur à celui de Gordon *et al.* qui n'ont détecté des *E. coli* que chez 5% des mammifères sauvages (Gordon et Cawling, 2003). Dans notre étude, la comparaison des fréquences de détections obtenues par les 2 techniques, moléculaire et phénotypique, confirme la sensibilité supérieure de la première (Tableau VI).

Espèces	Nombre d'échantillons	Prévalence par qPCR	Prévalence par étalement
Sangliers	15	100%	93%
Cerfs	17	100%	65%
Lapins sauvages	15	100%	100%
Cochons	15	100%	93%
Chevaux	15	100%	93%
Moutons	15	100%	73%
Vaches	15	100%	100%
Lapins domestiques	15	66%	0%
Poulets	15	100%	100%
Humains (*)	15	100%	100%

(*) 15 sujets sélectionnés au hasard issus de l'étude Smati *et al* 2013.

Tableau VI : Prévalence de *E. coli* dans les selles d'animaux domestiques, sauvages et d'humains.

Sur le plan quantitatif, nous retrouvons des taux de portage plus importants chez l'homme. Chez l'animal, ces taux sont variables en fonction de l'espèce, du régime et du mode de vie (Tableau 1 de l'article 2). Les espèces omnivores ont des taux plus élevés que les herbivores. Cela pourrait s'expliquer par la composition des selles beaucoup plus riches en fibres chez les

seconds que chez les premiers, mais également par la variété du régime des premiers qui conditionne des apports de souches d'origine alimentaire plus fréquents et plus importants.

Si l'on considère les groupes phylogénétiques, une concordance a été mise en évidence chez les animaux entre le groupe phylogénétique dominant et celui obtenu par typage d'une souche prise au hasard par individu. Cela montre que l'hypothèse du clone dominant validé pour l'homme lors du premier article est confirmée chez les animaux étudiés. Cette constatation permet de valider les comparaisons de fréquence des groupes dominants avec les résultats antérieurs de la littérature.

Les animaux sauvages se sont différenciés des animaux domestiques par de plus faibles quantités de *E. coli* totaux et par un nombre moins élevé de groupes phylogénétiques détectés par animal.

Ce plus faible portage, moins polymorphe, pourrait s'expliquer par des raisons épidémiologiques.

En effet, la densité de population est beaucoup plus faible chez les animaux sauvages prélevés. Ainsi, dans la forêt de Fontainebleau, il y a entre 500 et 700 cerfs pour 17 000 hectares ce qui correspond à une densité d'environ 4 individus pour 100 hectares. Les sangliers sont entre 1300 et 1500 pour la même superficie soit une densité d'environ 8 individus pour 100 hectares (JM Cacouault, Office National des Forêts, communication personnelle).

Les animaux domestiques qui ont été prélevés lors de notre travail vivent dans des conditions de promiscuité incomparables : 50 vaches partagent une étable de quelques centaines de mètres carrés, 500 moutons une bergerie de la même surface, les porcs sont parqués à 50 à 100 individus dans 100 m² et les poulets à 1000 individus dans la même surface. Ils reçoivent également une nourriture industrielle, des vaccins et des traitements antibiotiques.

Ces pressions de sélection se traduisent par des fréquences élevées de résistance aux antibiotiques. Les souches des animaux sauvages sont sensibles alors que celles des animaux domestiques présentent des taux appréciables de résistance à l'amoxicilline, à l'amoxicilline-acide clavulanique et au cotrimoxazole. Ces résistances ont été déterminées sur l'antibiogramme du clone dominant. En parallèle, les selles ont été étalées sur géloses ESBL (BioMérieux®) pour la détection de BLSE. Par cette méthode, un *E. coli* producteur de BLSE a été détecté chez 26% des vaches, ce taux atteint 100% chez les poulets (Figure 15).

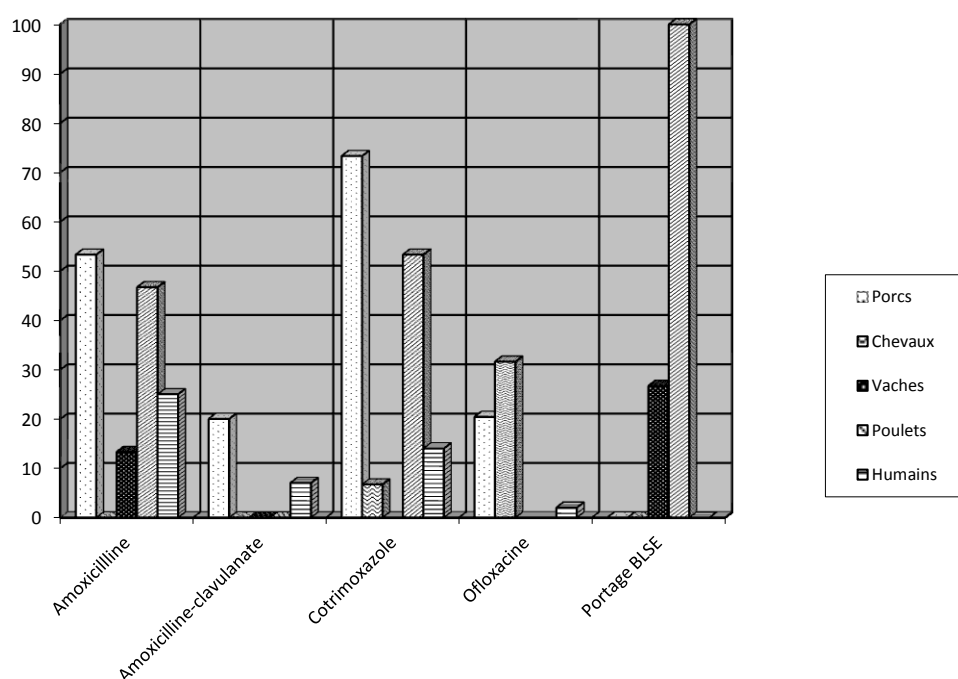


Figure 15 : Représentation graphique des taux de résistance des flores commensales des animaux domestiques et des humains. Aucune résistance à l'amikacine n'a été retrouvée. Aucune résistance n'a été retrouvée chez les animaux sauvages.

La caractérisation des groupes phylogénétiques confirme les données de la littérature en montrant que les groupes B1, B2 et A sont retrouvés comme groupes phylogénétiques majoritaires.

Ainsi 3 entérocolotypes sont caractérisables et répartis chez les animaux en fonction de plusieurs spécificités :

- l'entérocolotype B2 est mis en évidence chez les herbivores sauvages. Il est, comme chez l'homme, dominant et accompagné de peu d'autres groupes phylogénétiques minoritaires. Mais, à la différence du modèle humain, il n'est jamais exclusif.
- l'entérocolotype B1 est retrouvé chez les herbivores domestiques ce qui est conforme aux données de la littérature. Le groupe B1 est dominant mais accompagné de groupes minoritaires détectables.
- l'entérocolotype A est retrouvé chez les omnivores sauvages et domestiques. Comme chez l'homme, le groupe A dominant est accompagné de groupes minoritaires détectables.

La mise en évidence d'une fréquence importante de souches de type B2 commensales chez l'homme et les animaux sauvages et sa rareté chez les animaux domestiques présente un problème d'interprétation.

En effet, la plus grande proximité de l'homme avec les animaux domestiques pourrait conduire à un résultat inverse. D'autre part, le caractère non exclusif du groupe phylogénétique B2 chez l'animal, alors qu'il est retrouvé exclusif chez 21% des hommes, nous a amené à émettre l'hypothèse d'une différence de nature entre les souches B2 humaines et animales.

Ainsi, des distinctions ont été mises en évidence entre les phylums en fonction de leurs origines. Les sous-groupes de B2 ainsi que les types O décrits jusqu'alors ont caractérisés les

souches humaines alors que les souches animales étaient fréquemment non-typables en terme de sous-groupe B2 et de type O. Cela indique qu'elles appartiennent à des clones non encore répertoriés de l'espèce ; les schémas de typage ayant été principalement constitués avec des souches d'origine humaine. Un résultat comparable a été obtenu avec les souches B2 d'origine animale isolées en Guyane (Lescat *et al.*, 2013).

De plus, les souches d'origine animale ont présenté significativement moins de facteurs de virulence que les souches d'origine humaine et 2 gènes de facteurs de virulence, *papC* et *usp* ont été significativement moins fréquemment détectés chez les souches animales.

Cela confirme le caractère très adapté et super colonisateur des souches B2 chez l'homme et explique leur présence en situation de groupe phylogénétique exclusif. Par contre, la production de phages/colicine n'a pas été différente entre les souches humaines et animales.

Pour les souches d'origine humaine, d'autres gènes de facteurs de virulence comme *sfa/foc* et *clbQ* étaient significativement plus fréquents parmi les souches correspondant au groupe phylogénétique B2 en situation exclusive qu'en situation seulement dominante.

Une certaine hiérarchie peut donc être mise en évidence entre les souches B2 humaines exclusives, humaines dominantes et animales dominantes.

Dans l'hypothèse déjà évoquée où le portage finalement banal de ces souches B2 suréquipées ne donnerait pas lieu chez les porteurs à une fréquence ou une gravité plus particulière des infections extra intestinales à *E. coli*, cela indiquerait que les facteurs de virulence intrinsèques ne sont, chez l'homme sain, que des facteurs de colonisation et que l'existence d'un état morbide ne dépend finalement presque exclusivement que de la défaillance de l'hôte.

L'ensemble de ces résultats montre que l'approche quantitative par qPCR de l'estimation des proportions des groupes phylogénétique permet une évaluation réaliste de la diversité de l'espèce dans une niche écologique donnée et qu'elle rend bien compte du comportement des souches et de leur capacité à coloniser cette niche.

La même stratégie pourrait être appliquée avec profit à d'autres situations physiologiques ou physiopathologiques.

**IMPACT DU HPI (ILOT DE HAUTE
PATHOGENICITE OU PAIIV) SUR
LA VIRULENCE *D'ESCHERICHIA
COLI* EN FONCTION DU FOND
GENETIQUE**

I. Introduction

Le fer est essentiel à la croissance bactérienne, ce qui fait des systèmes de captation du fer ou sidérophores un facteur de commensalisme indispensable à la survie de la bactérie et à son maintien dans sa niche naturelle en entrant en compétition avec d'autres microorganismes pour l'acquisition de cet élément. Dans les hôtes mammifères, la quantité de fer libre est limitée et le fer se retrouve lié ou complexé à des protéines, les bactéries étant ainsi dans un milieu pauvre en fer (Wandersman et Delepelaire 2004).

Or la séquestration du fer vis-à-vis des organismes infectieux est aussi l'une des défenses innées contre l'infection (Miller *et al.*, 2006). Le fait qu'un microorganisme acquiert des sidérophores supplémentaires peut donc lui permettre de résister à cet environnement, le sidérophore devenant de fait un facteur de virulence. *E. coli* étant une bactérie modèle de transition entre les deux états commensalisme et virulence, l'étude des sidérophores dans différents modèles de virulence expérimentale et dans différents fonds génétiques peut nous apporter une meilleure compréhension de cette dualité.

Les souches d'ExPEC produisent jusqu'à 4 sidérophores : entérobactine, salmochéline, aërobactine et yersiniabactine (Braun *et al.*, 2002). Ces sidérophores sont fortement surexprimés *in-vivo* dans des modèles d'UTI (infections urinaires) (Snyder *et al* 2004).

Le HPI précédemment décrit est le PAI responsable de la sécrétion du sidérophore yersiniabactine.

On s'intéressera dans notre étude à l'effet de la mutation des principaux gènes portés par le HPI: *irp2*, gène de synthèse de la yersiniabactine, *fyuA* qui code pour le récepteur de la yersiniabactine (et également impliqué dans la synthèse du biofilm, Hancock *et al* 2008) et enfin *ybtA* régulateur du HPI et régulateur pléiotrope.

Le HPI est retrouvé dans des souches virulentes de *E. coli* en particulier celles du groupe phylogénétique B2. Il est d'arrivée récente dans l'espèce et présente une grande homologie

avec celui de *Yersinia pestis* (99%, Rakin et Heesemann 1995). Ce même îlot génomique peut être nécessaire à la grande virulence d'une espèce telle que *Yersinia pestis*, participer à la virulence de souches d'une autre espèce sans pour autant être nécessaire (*E. coli* entéro-agrégatif) ou encore être présent parmi les souches commensales d'un individu (*E. coli*).

De plus, ce HPI pouvant être fonctionnel ou non (selon l'intégrité du récepteur FyuA), son maintien dans l'espèce pose aussi la question de son rôle en dehors de la captation du fer. Cette fonction qui reste à définir pourrait être médiée par le sidérophore produit en faibles quantités et qui se comporterait comme molécule de signalisation et/ou par YbtA qui semble être un régulateur pleiotrope. Cet îlot génomique pourrait contribuer à augmenter la fitness de la souche, sa flexibilité métabolique ou augmenter son pouvoir pathogène en fonction du fond génétique sur lequel cet îlot arrive.

Les études sur *K. pneumoniae* montrent également que selon le modèle de virulence, et en particulier la voie d'inoculation (instillation intra-nasale *versus* injection intrapéritonéale), le rôle du HPI peut-être ou non critique dans la virulence.

Chez *E. coli*, on retrouve cette même hétérogénéité selon les souches et les modèles: le HPI est important dans la phase bactériémique dans un modèle de septicémie chez la souris pour les souches IAI51 et IAI52 (Schubert *et al.*, 2002), mais n'a pas d'impact dans un modèle de méningite néonatale chez le rat nouveau-né (souche C5) (Nègre *et al.*, 2004), ni dans un modèle d'infection urinaire ascendante chez la souris (souche 536) (Brzuszkiewicz 2006), ni dans un modèle de septicémie chez la souris (souche 536) (Tourret *et al.*, 2010) .

Afin de tester l'influence du fond génétique sur la virulence médiée par le HPI, trois souches de *E. coli* B2 ont été choisies, ce sont tous des isolats naturels uropathogènes et de ST (séquence type) différents : CFT073, 536 et NU14 respectivement des sous-groupes II (STc73), III (STc127) et IX (STc95) et de composition variable en sidérophores.

Ces souches représentent donc les principaux sous-groupes de B2 qui regroupent les souches ExPEC. Ainsi, les sous-groupes II et IX représentent plus de 50% des isolats de B2 ExPEC.

Des mutants des trois principaux gènes du HPI ont été construits par l'équipe du Pr. Sören Schubert de l'Institut Max Von Pettenkoffer de Munich et recomplémentés sur le chromosome ou par un plasmide.

Leur croissance a été étudiée, et leur virulence testée dans 2 modèles de pathologie expérimentale : d'une part la septicémie de la souris et d'autre part la résistance à la phagocytose par l'amibe sociale *D. discoideum*.

De plus, des études de survie intra-amibes ont été entreprises (expériences d'invasion cellulaire) avec dénombrement des bactéries intra-amibes et des amibes. L'effet du surnageant des souches sauvages et de leur mutants sur la croissance de l'amibe a été observé.

Enfin, la souche NU14 appartenant au sous-groupe IX, et étant très proche des souches de méningites néonatales qui portent toutes le HPI, nous avons testé d'autres mutants d'isolats naturels de ce sous-groupe (C5 et S88) et leurs mutants.

Une partie de ces travaux consacrée à la virulence chez l'amibe a été effectuée en collaboration avec le Dr. Sandrine Adiba.

II. Matériel et méthodes

1). Matériel

a). Souches bactériennes:

Les souches de *E. coli* 536 et CFT073 et NU14 sont trois souches isolées d'infection urinaire devenues des modèles pour l'étude génétique des PAIs et de leurs rôles dans la virulence extra-intestinale.

Ces souches appartiennent toutes au groupe phylogénétique B2, et ont été choisies car appartenant à trois sous-groupes différents, les sous-groupes II, III et IX qui regroupent majoritairement les souches d'ExPEC. Il est à noter qu'ils font partie des quatre sous-groupes (II, III, VIII, and IX) qui ont le moins de diversité intra-spécifique (moins de 0.1%) (Le Gall *et al.*, 2007). Par la suite, d'autres isolats du sous-groupe IX de B2 (souches de méningites néonatales C5 et S88) ont été testés.

Souche 536

La souche pathogène UPEC (*E. coli* uropathogène) 536 (O6:K 15:H31) appartient au sous-groupe III du groupe phylogénétique B2. Cette souche a été isolée du tractus urinaire d'un patient atteint d'une pyélonéphrite (Hacker *et al.* 1983). Son génome est entièrement séquencé, il mesure 4938 kb, et ne contient qu'une seule région prophagique. Une étude génomique (Hacker et Kaper, 2000) a mis en évidence sept PAIs dont le HPI. Il s'agit d'une souche résistante au sérum humain, à la bile ovine et au lysozyme. Elle possède des enzymes impliquées dans l'acquisition et la capture du fer. Elle est très motile et sécrète de l'hémolysine (Diard *et al.*, 2007). C'est une souche très virulente chez la souris (Picard *et al.*, 1999) ainsi que chez le nématode *Caenorhabditis elegans* (Diard *et al.*, 2007).

Le rôle du HPI dans la virulence de 536 est incertain. En effet, aucune diminution de virulence n'a été notée lorsqu'un mutant de 536 délété de cet îlot (536 Δ HPI) a été testé dans un modèle d'infection urinaire (Brzuszkiewicz *et al.* 2006). De même, ce mutant de 536 n'était pas affecté dans sa virulence dans le modèle de nématode *C. elegans*. Le mutant 536 Δ HPI dans l'étude de Médéric Diard (2010) ne se maintenait pas dans un modèle de colonisation intestinale chez la souris quand il était en compétition avec la souche sauvage. Il utilisait pourtant de la même façon le gluconate pour sa croissance, principale source de carbone dans l'intestin. Il résistait aussi bien que la souche sauvage à différents stress biologiques. Il n'était pas affecté dans sa virulence dans un modèle de septicémie chez la souris (Tourret *et al.*, 2010).

Cette souche produit les sidérophores entérobactine, salmochéline et yersiniabactine mais ne produit pas l'aérobactine.

Souche CFT073

La souche CFT073 est une souche isolée des urines et du sang d'une patiente ayant présenté une septicémie à point de départ urinaire (Mobley *et al.* 1990). Elle appartient au sous-groupe II du groupe phylogénétique B2. Son génome est entièrement séquencé.

Elle dérive du groupe clonal O6:K2:H1, dont sont issues de nombreuses souches agents de pyélonéphrites et d'urosepsis (Orskov 1982, Johnson et Stell 2000). Ses PAIs ont été comparés à ceux de la souche 536 (Figure 16). Les PAI I–III et V de 536 contiennent des gènes absents de CFT073 (telles que des gènes d'adhésines) ou des gènes qui sont localisés dans des îlots similaires localisés ailleurs dans le génome de CFT073. Par contre, les HPI sont très semblables (Brzuszkiewicz *et al.*, 2006).

CFT073 possède les facteurs de virulence suivants: *neuC*, *kpsE*, *sfa/foc*, *iroN*, *aer*, *iha*, *papC*, *papG(II)*, *hlyC*, *sat*, *ireA*, *usp*, *ompT*, *malX*, *fyuA*, and *irp2*. Elle n'a pas de résistance acquise aux antibiotiques.

CFT073 comporte le HPI en association avec l'ARN_t *asnT*, mais comporte une IS (séquence d'insertion) dans le gène *irp2*.

Elle produit l'entérobactine, la salmochéline et l'aérobactine (Welch *et al.*, 2002). Elle n'est pas capable d'utiliser la yersiniabactine comme sidérophore, mais elle comporte les gènes du HPI dont le régulateur YbtA, et peut rendre compte de l'effet de ce dernier en tant que régulateur pleiotrope en dehors de sa fonction d'induction de sécrétion de sidérophores.

Figure 16: Distribution des PAIs portant les facteurs de virulence (FV) majeurs dans les souches 536 et CFT073. Les ilots ont été indiqués selon leur insertion chromosomique près des loci ARNt. Quelques FV sont indiqués selon leur localisation sur les PAIs (Brzuszkiewicz 2006).

Souche NU14 :

La souche NU14 a été isolée d'une cystite. Elle a joué un rôle dans la découverte et la dissection du mécanisme moléculaire de FimH (adhésine fimbriale de type 1), médiant l'adhérence et l'invasion des cellules urothéliales (Mulvey *et al.*, 1998).

Elle dérive du même sous-clone de *E. coli* O18:K1:H7 que la souche archétypale de méningite néonatale RS218, et porte les mêmes facteurs de virulence (Jonhson Sokurenko *et al.*, 2001). Elle appartient au sous-groupe IX du groupe phylogénétique B2.

Elle sécrète l'entérobactine, la salmochéline et la yersiniabactine mais pas l'aérobactine (Tableau VII).

Tableau VII : Caractéristiques de la souche NU14 et comparaison avec les souches de *E. coli* RS218, CFT073 et 536 (d'après Johnson *et al.*, 2001 et Adiba *et al.*, 2010).

Gènes de virulence												
Souches	K1	<i>chuA</i>	<i>sfa/foc</i>	<i>iroN</i>	<i>aer</i>	<i>papC</i>	<i>papG</i>	<i>hly</i>	<i>cnfI</i>	<i>hra</i>	<i>fyuA</i>	<i>irp2</i>
RS218	+	+	+	+	-	+	III	+	+	+	+	+
536	-	+	+	+	-	+	III	+	-	+	+	+
NU14	+	+	+	+	ND	+	III	+	+	ND	+	+
CFT073	-	+	+	+	+	+	II	+	-	-	+	+

Caractères phénotypiques						
Souches	Résistant	Résistant	Résistant	motilité	Septicémie	Amibe
	sérum	lysozyme	bile		souris	GR
RS218	+	+	+	+	K	GR
536	+	+	+	+	K	GR
NU14	ND	ND	ND	ND	ND	ND
CFT073	+	+	-	+	K	GR

ND : non déterminé, GR : grazing resistant, résistant au broutage par l'amibe, K : « killer », tuant 9 à 10 souris sur 10 dans le modèle de septicémie (Picard *et al.*, 1999).

Souche B REL606 :

C'est une souche de *E. coli*, avirulente et appartenant au phylogroupe A. Elle a été beaucoup utilisée dans l'étude de l'évolution (souche de Richard Lenski, Michigan State University). Elle a été isolée en 1918 par Félix d'Hérelle à partir de fèces d'homme sain.

Elle sert de contrôle négatif dans les modèles de virulence testés.

Souche C5 :

C'est une souche de méningite néonatale (Bortolussi *et al.*, 1978), elle est représentative du groupe clonal américain majeur O18:K1:H7 et possède le HPI, avec les gènes *iroN* et *chuA*, mais ne produit pas d'aérobactine (Bonacorsi *et al.*, 2000, Bonacorsi *et al.*, 2003). Elle appartient au sous-groupe IX des B2 ou ST95 d'Achtman.

Souche S88 :

C'est une souche représentative du groupe clonal français O45:K1:H7 qui représente un tiers des méningites néonatales en France (Peigne *et al.*, 2009). Cette souche a été isolée en 1989 du liquide céphalo-rachidien d'un nouveau-né. Cette souche porte un large plasmide pS88 de 133,853 pb codant pour 144 protéines et qui porte trois différents systèmes de sidérophores (aérobactine, salmochéline et les gènes *sitABCD*) et d'autres gènes de virulence putatifs (*iss*, *etsABC*, *ompTP*, et *hlyF*) (Bonacorsi *et al.*, 2003, Peigne *et al.*, 2009) (Figure 17). Il est à noter que des copies des gènes *sitABCD* existent sur le chromosome. Les autres gènes chromosomiques sont l'antigène K1, les P fimbriae avec l'adhésine PapGII, et le HPI.

L'antigène capsulaire K1 et le sidérophore salmochéline semblent être les seuls traits de virulence spécifiques partagés par ce clone et le groupe clonal O18:K1:H7 (type souche C5), même si ces deux groupes clonaux appartiennent tous deux au même sous-groupe virulent IX de B2 (Bidet *et al.*, 2007, Le Gall *et al.*, 2007).

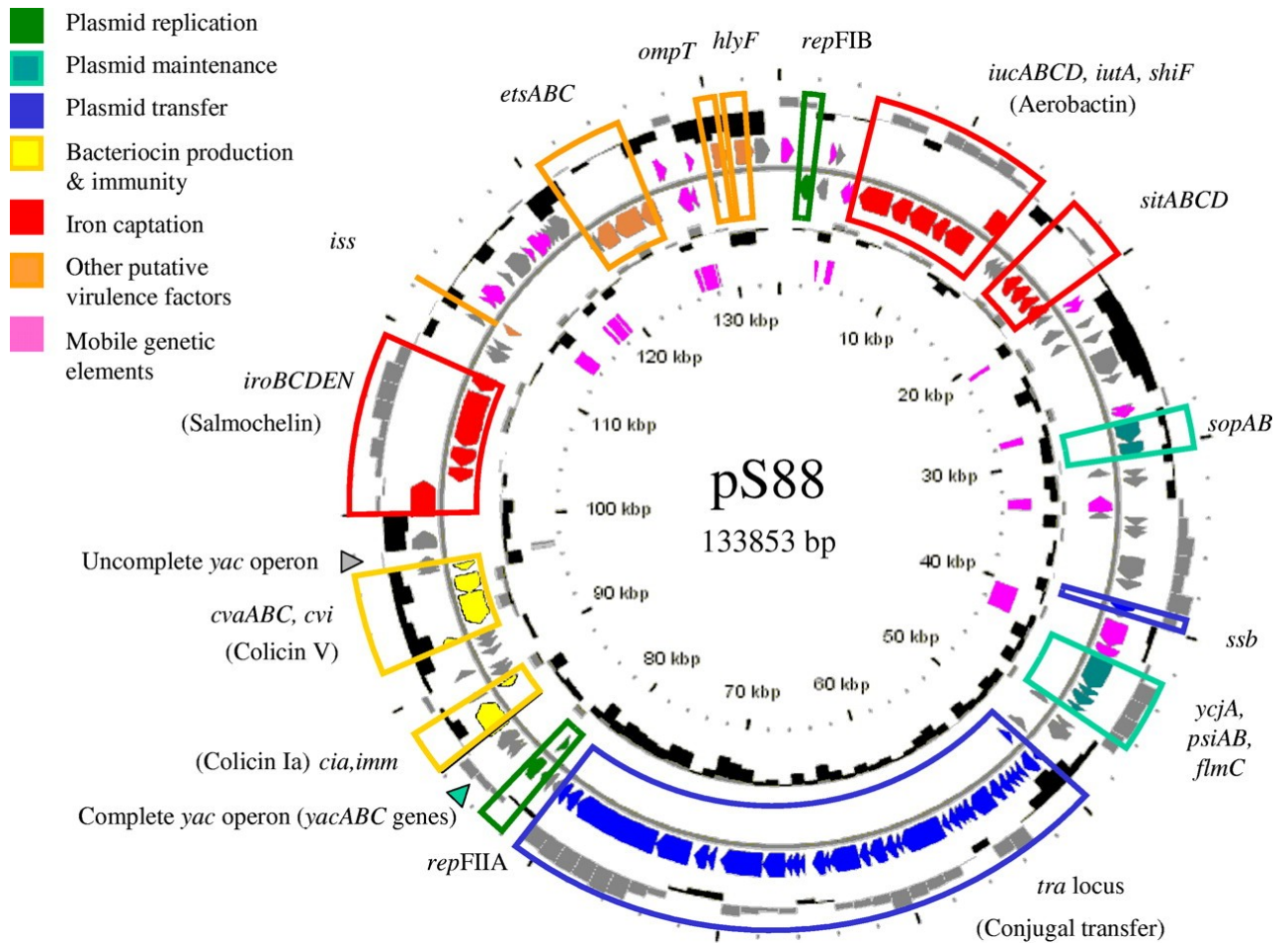


Figure 17: Représentation circulaire du plasmide de la souche de *E. coli* S88 (pS88)
(Peigne *et al.*, 2009).

b). Plasmides :

pS88cat : plasmide pS88 rendu résistant au chloramphénicol (Cm) (Peigne *et al.*, 2009).

pWSK30: plasmide à faible nombre de copies, résistant à l'ampicilline (Ap), possédant une origine de répllication pSC101 (Wang et Kushner 1991).

pACYC184: plasmide à faible nombre de copies, résistant aux tétracyclines (Tet) contenant le système de répllication du miniplasmide p15A (Rose 1988)

pCp1: vecteur suicide s'intègre dans le chromosome, porte la région entière du HPI de 35 kb codant pour le système yersiniabactine (*irpI-9*, *fyuA*, *ybtA* et 400 nucléotides de *intB*), chloramphénicol résistant (Schubert *et al.*, 2002, Pelludat *et al.*, 2002).

pCR-XL Topo : plasmide d'Invitrogen®, à haut nombre de copies, porte un promoteur de type T7, résistant à la kanamycine (Kn) et à la zéocine (Zn).

Tableau X : Récapitulatif des souches et plasmides utilisés dans cette étude

Souches	Plasmides	Résistances	Origine
CFT073			
536			
NU14			
B REL606			
S88			S. Bonacorsi
C5			
CFT073 <i>ybtA</i> ::Kn		Kn	S. Schubert
CFT073 HPI ::Cm		Cm	S. Schubert
536 <i>ybtA</i> ::Kn		Kn	S. Schubert
536 <i>ybtA</i> ::Kn	pWKS30- <i>ybtA</i>	Kn, Ap	S. Schubert
536 <i>fyuA</i> ::Cm		Cm	S. Schubert
536 <i>fyuA</i> ::Cm	pACYC184- <i>fyuA</i>	Cm, Tet	S. Schubert
536 <i>irp2</i> ::Kn		Kn	S. Schubert
536 <i>irp2</i> ::Kn	pCRXLTopo- <i>irp2</i>	Kn, Zn	S. Schubert
536 ΔPAI III (salmochéline -)			Diard 2010
536 ΔHPI (yersiniabactine -)			Diard 2010
536 ΔPAI III-HPI (salmocheline et yersiniabactine -)			Diard 2010
536 ΔPAI I à VII (déléte des 7 PAIs)			Diard 2010
NU14 <i>ybtA</i> ::Kn 1		Kn	S. Schubert
NU14 <i>ybtA</i> ::Kn	pWKS30- <i>ybtA</i>	Ap, Kn	S. Schubert
NU14 <i>ybtA</i> ::Kn pCP1		Kn, Cm	S. Schubert
NU14 <i>ybtA</i> ::Kn 2 (deuxième mutant)		Kn	S. Schubert
NU14 <i>ybtA</i> ::Kn 2	pWKS30- <i>ybtA</i>	Ap, Kn	S. Schubert
NU14 <i>irp2</i> ::Kn		Kn	S. Schubert
NU14 Δ <i>irp2</i>	pWKS30- <i>irp2</i>	Ap	S. Schubert
NU14 <i>irp2</i> ::Cm (deuxième mutant)		Cm	S. Schubert
NU14 <i>irp2</i> ::Cm	pWKS30- <i>irp2</i>	Cm, Ap	S. Schubert
NU14 <i>fyuA</i> ::Cm		Cm	S. Schubert
NU14 <i>fyuA</i> ::Cm	pACYC184- <i>fyuA</i>	Cm, Tet	S. Schubert
NU14 HPI ::Kn			S. Schubert
C5 ΔHPI :: <i>cat-1</i>		Cm	S. Bonacorsi
C5 <i>fyuA</i> :: <i>cat-1</i>		Cm	S. Bonacorsi
C5 <i>fyuA</i> :: <i>cat-1</i>	pACYC184- <i>fyuA</i>	Cm, Tet	Cette étude
S88 ΔHPI :: <i>cat-1</i>		Cm	S. Bonacorsi
S88 <i>fyuA</i> :: <i>cat-1</i>		Cm	S. Bonacorsi
S88 ΔpS88			S. Bonacorsi

S88 Δ HPI :: <i>cat-1</i> Δ pS88		Cm	Cette étude
S88 <i>fyuA</i> :: <i>cat-1</i> Δ pS88		Cm	Cette étude
S88 <i>fyuA</i> :: <i>cat-1</i> Δ pS88	pACYC184- <i>fyuA</i>	Cm, Tet	Cette étude

c). Souches d'amibes :

Nous utilisons l'amibe sociale *D. discoideum* AX3 pour l'ensemble de nos expérimentations. Il s'agit d'une souche dite axénique capable de croître en utilisant les nutriments du milieu de culture. Comme décrit précédemment, cette amibe a déjà été utilisée pour mesurer la virulence de diverses souches bactériennes. Ce modèle de virulence permet de tester un grand nombre de souches (Froquet *et al.*, 2008) et à moindre coût.

Cette amibe croît dans un milieu dit HL5 et sa température de croissance optimale se situe entre 20 et 25°C.

Les amibes sont conservées à -80°C dans 500µL de HL5 additionné de 500 µL de diméthylsulfoxyde (DMSO).

d). Souris :

Des souris OF1 de 14-16 grammes femelles de Charles River® (L'Arbresle, France) ont été utilisées.

e). Milieus et antibiotiques:

Antibiotiques : ampicilline (100 µg/ml), kanamycine (25 µg/ml), chloramphénicol (20 µg/ml), tétracycline (15 µg/ml), zéocine (25 µg/ml).

HL5 (milieu de culture des amibes) :

Protéose peptone 5 g, thiotone E peptone 5 g, glucose 10 g, extrait de levure 5 g, Na₂HPO₄ 0.35 g, KH₂PO₄ 0.35 g, agar 15 g. H₂O distillée qsp 1L. Après ajustement du pH à 6,5 +/- 0,1, le milieu est autoclavé.

Tampon « Magnésium, Calcium, Phosphate Buffer » MCPB (tampon de développement) :

Na₂HPO₄ 1,42 g ; KH₂PO₄ 1,36 g, MgCl₂ 0,19 g ; CaCl₂ 0,03 g dans 1 l d'eau pure à pH 6,5 +/- 0,1. La solution est filtrée.

2). Méthodes :

a). Déterminants bactériens :

La séquence complète du génome de la souche NU14 n'étant pas encore disponible, la présence de gènes de virulence caractéristiques des ExPEC a été recherchée par PCR et comprend : des adhésines (*papC*, *papG*, incluant les allèles de *papG* qui sont *papGII* et *papGIII* ; *sfa/foc*, *hra*, *iha* et *ibeA*), des toxines (*hlyC*, *cnfI*, *sat*, *usp* et *clbQ*), des systèmes de captation du fer (*fyuA*, *irp2*, *iroN*, *iucC* et *ireA*), des protectines (*neuC*, *ompT* et *TraT*), ainsi qu'un gène codant pour une protéine uropathogène spécifique *usp* (comme décrit précédemment dans Lefort *et al.*, 2011, Johnson et Clabots *et al.*, 2008) (Tableau XI).

Tableau XI : Séquences des amorces utilisées pour détecter les gènes de virulence extraintestinale.

Amorces	Séquence (5'-3')	Gène	Longueur du produit de PCR (pb)	Température d'hybridation (°C)
papC.1	GACGGCTGTACTGCAGGGTGTGGCG	<i>papC</i>	328	60
papC.2	ATATCCTTTCTGCAGGGATGCAATA			
papG.II.1	GGGATGAGCGGGCCTTTGAT	<i>papG</i>	190	63
papG.II.2	CGGGCCCCCAAGTAACTCG			
papG.III.1	GGCCTGCAATGGATTTACCTGG		258	63
papG.III.2	CCACCAAATGACCATGCCAGAC			
aer.1	AAACCTGGTTTACGCAACTGT	<i>iucC</i>	269	59
aer.2	ACCCGTCTGCAAATCATGGAT			
hlyC.1	AGGTTCTTGGGCATGTATCCT	<i>hlyC</i>	556	55
hlyC.2	TTGCTTGCAGACTGCAGTGT			
iroN.1	AATCTTACCGCAATAGCGC	<i>iroN</i>	150	60
iroN.2	GCTATAACGCACGGAGTTAC			
traT.1	GGTGTGGTGCGATGAGCACAG	<i>traT</i>	290	59
traT.1	CACGGTTCAGCCATCCCTGAG			
fyuA.1	TGATTAACCCCGCGACGGGAA	<i>fyuA</i>	780	55
fuyA.2	CGCAGTAGGCACGATGTTGTA			
irp2.1	AAGGATTCGCTGTTACCGGAC	<i>irp2</i>	280	60
irp2.2	TCGTCGGGCAGCGTTTCTTCT			
sfa.1	CTCCGAGAACTGGGTGCATCTTAC	<i>sfa/foc</i>	410	60
sfa.2	CGGAGGAGTAATTACAAACCTGGCA			
kpsE.1	AGTAGAGTTCAACACCGAGC	<i>kpsE</i>	205	59
kpsE.12	CAGCAACAGTTTCAGGGTGC			
cnf.1	CAGTGACCGGATCTCCGTAT	<i>cnfI</i>	240	60
cnf.2	CGTGTAATTCTTCTGTACTTCC			
hra.1	CAGAAAACAACCGGTATCAG	<i>hra</i>	260	60
hra.2	ACCAAGCATGATGTCATGAC			

Amorces	Séquence (5'-3')	Gène	Longueur du produit de PCR (pb)	Température d'hybridation (°C)
ibeA.1	TTACCGCCGTTGATGTTATCA	<i>ibeA</i>	171	63
ibeA.2	CATTAGCTCTCGGTTACGCT			
iha.1	CTGGCGGAGGCTCTGAGATCA	<i>iha</i>	829	55
iha.2	TCCTTAAGCTCCCGCGGCTGA			
Sat.1	GCAGCTACCGCAATAGGAGGT	<i>sat</i>	937	55
Sat.2	CATTCAAGTACCGGGGCCCTA			
ireA.1	GATGACTCAGCCACGGGTAA	<i>ireA</i>	254	60
ireA.2	CCAGGACTCACCTCACGAAT			
usp.1	ACATTCACGGCAAGCCTCAG	<i>usp</i>	440	60
usp.2	AGCGAGTTCCTGGTGAAAGC			
ompT.1	ATCTAGCCGAAGAAGGAGGC	<i>ompT</i>	559	55
ompT.2	CCCGGGTCATAGTGTTCATC			
NeuC.1	GTGAAAAGCCTGGTAGTGTG	<i>neuC</i> (K1)	255	60
NeuC.2	AACCAGTGTACGCGTACTG			
clbQ.1	CTTGTATAGTTACAACTATTTC	<i>pks</i>		55
clbQ.2	TTATCCTGTTAGCTTTCGTTTC			

b). Vérification des résistances aux antibiotiques

Pour les souches recomplémentées par un plasmide, la pression de sélection était maintenue tout au long des processus de culture. Les clones ont été inoculés dans du bouillon « Lysogeny Broth » (LB) additionné d'antibiotique marqueur. Après incubation une nuit à 37°C, les cultures étaient ensemencées pures sur géloses LB additionnées d'antibiotiques afin de vérifier la présence du plasmide.

Il a également été possible de maintenir cette pression de sélection dans le modèle amibe par adjonction d'antibiotiques dans les géloses de HL5. Il a été vérifié en parallèle que cela n'affectait pas la culture amibienne.

c). Obtention des mutants et recomplémentation plasmidique :

Certains mutants ont été construits par l'équipe de Sören Schubert (Tableau X). Les mutants recomplémentés avaient été vérifiés dans un bioessai (Schubert 2000) sur la base de leur capacité à produire le sidérophore (Giuseppe Magistro, Sören Schubert communication personnelle). Ces mutants ont été construits par méthode de Datsenko-Wanner (Datsenko et Wanner 2000). De plus certains mutants (mutants *ybtA* de NU14, mutants *irp2* de NU14) ont été construits sur des jours différents afin de s'assurer de l'absence de « side-effect » ou d'effet collatéral indésirable sur l'ADN chromosomique des souches. Les mutants des souches C5 et S88 ont été construits par l'équipe de Stéphane Bonacorsi. Les mutants C5 et S88 délétés de *fyuA* ont été recomplémentés par le plasmide pACYC184-*fyuA* au cours de cette étude, ceci a nécessité l'extraction de plasmides, la préparation des cellules électrocompétentes et leur transformation par électroporation.

Après une préculture d'une nuit dans 10 ml de LB de la souche à transformer, celle-ci est ensemencée dans 100 ml de LB et incubée environ 3h à 37°C sous agitation jusqu'à l'obtention d'une densité optique de 0,5 à 600 nm (phase exponentielle de croissance).

Après, suivent des étapes de lavages-concentrations des cellules à 4°C, en reprenant les culots des 100 ml de culture bactérienne dans 50 ml d'une solution d'eau distillée avec 10 % de glycérol après centrifugation 5 mn à 7500 rpm, puis dans 25 ml, 10 ml, 5 ml et 1 ml jusqu'à reprise du culot final dans 150 µl de solution à 10% de glycérol. Les cellules étaient soumises à une électroporation immédiatement ou conservées à -80°C.

Pour effectuer l'électroporation, 50µl de cellules électrocompétentes étaient mélangées à 3 µl d'ADN de plasmide (obtenu après extraction par miniprep QIAGEN®) et placées dans la cuve de l'électroporateur, voltage à 1800 V, constante de temps 5,5 msec.

Après électroporation, les bactéries étaient rapidement remises en suspension dans du LB préchauffé à 37°C, incubées sous agitation pendant 2h, puisensemencées sur géloses LB additionnées d'antibiotiques en fonction du plasmide.

d). Cure du plasmide pS88

La souche S88 et ses mutants ont été cultivés 18h dans du milieu LB à 37°C sous agitation. La culture a alors été diluée à 10⁵ UFC/ml et des concentrations sérielles (2,5, 5, and 10%) de sodium dodécyl sulfate ajoutées. Après 18h de croissance sous agitation, les cultures ont étéensemencées sur des géloses LB, incubées pendant 18h et les colonies obtenues ont été testées pour leur capacités à produire des colicines en les repiquant sur une gélose LBensemencée avec une suspension d'une souche de *E. coli* K-12, sensible à la colicine. Les souches n'inhibant pas la croissance de K-12 avaient perdu le plasmide ColV. La perte du plasmide a été confirmée par des PCR ciblant les gènes de virulence portés par ce plasmide tels que ceux codant pour les sidérophores *iroN* et *iucC* (Peigne *et al.*, 2009).

e). Protocole de pathologie expérimentale avec *D. discoideum* :

Une semaine avant l'expérimentation, les amibes sont décongelées et mises en culture dans 10 ml de milieu HL5 dans des flacons de culture cellulaire de 25 ml et incubées à 23 °C pendant 7 jours.

Les bactéries sont cultivées la veille de l'expérimentation dans 10 ml de milieu HL5 (en BD Falcon de 25 ml) à 37°C sous agitation constante (200 rpm).

Les bactéries sont lavées par centrifugation 15 mn à 10.000 rpm puis le culot est repris dans 10 ml de tampon MCPB. Nous avons testé 3 à 4 quantités d'amibes : 10, 100, 1000 et 10 000 cellules. Trois réplicats étaient réalisés par souche, et toutes les expérimentations ont été faites deux fois.

Trois-cents µL de culture bactérienne lavées dans le MCPB étaient ensemencées sur des boîtes de Pétri de 55 mm de diamètre contenant de la gélose HL5. Les boîtes sont mises à sécher sous hotte pendant 30 mn.

Pendant le séchage des cultures bactériennes, les solutions d'amibes de concentrations connues sont préparées. La culture amibienne arrivée à confluence au 7^{ème} jour est lavée après centrifugation à 2000 rpm pendant 7 mn avec 10 ml de MCPB. Les amibes sont ensuite comptées en cellule Kovalslide® et ajustées à 3 concentrations correspondant à 10, 100 ou 1000 amibes pour 300 µL de suspension.

Les boîtes après étalement de 300 µL de suspension d'amibes sont à nouveau mises à sécher sous hotte puis incubés à 23°C scellées dans du parafilm pour éviter la dessiccation. Les boîtes étaient examinées à J3, J6, J9 pour détecter l'apparition de plages de lyse.

L'apparition de plages de lyse correspondait au « broutage » de la bactérie par l'amibe et définissait le phénotype dit Grazing sensitive (GS). La souche avirulente utilisée comme

témoin négatif était REL606 avec apparition de plages de lyse à 10, 100, 1000 amibes (Figure 18).

A l'inverse, les souches virulentes telles que 536 (536 était utilisée comme témoin positif) résistaient au broutage par l'amibe et définissait le phénotype Grazing résistant ou GR.

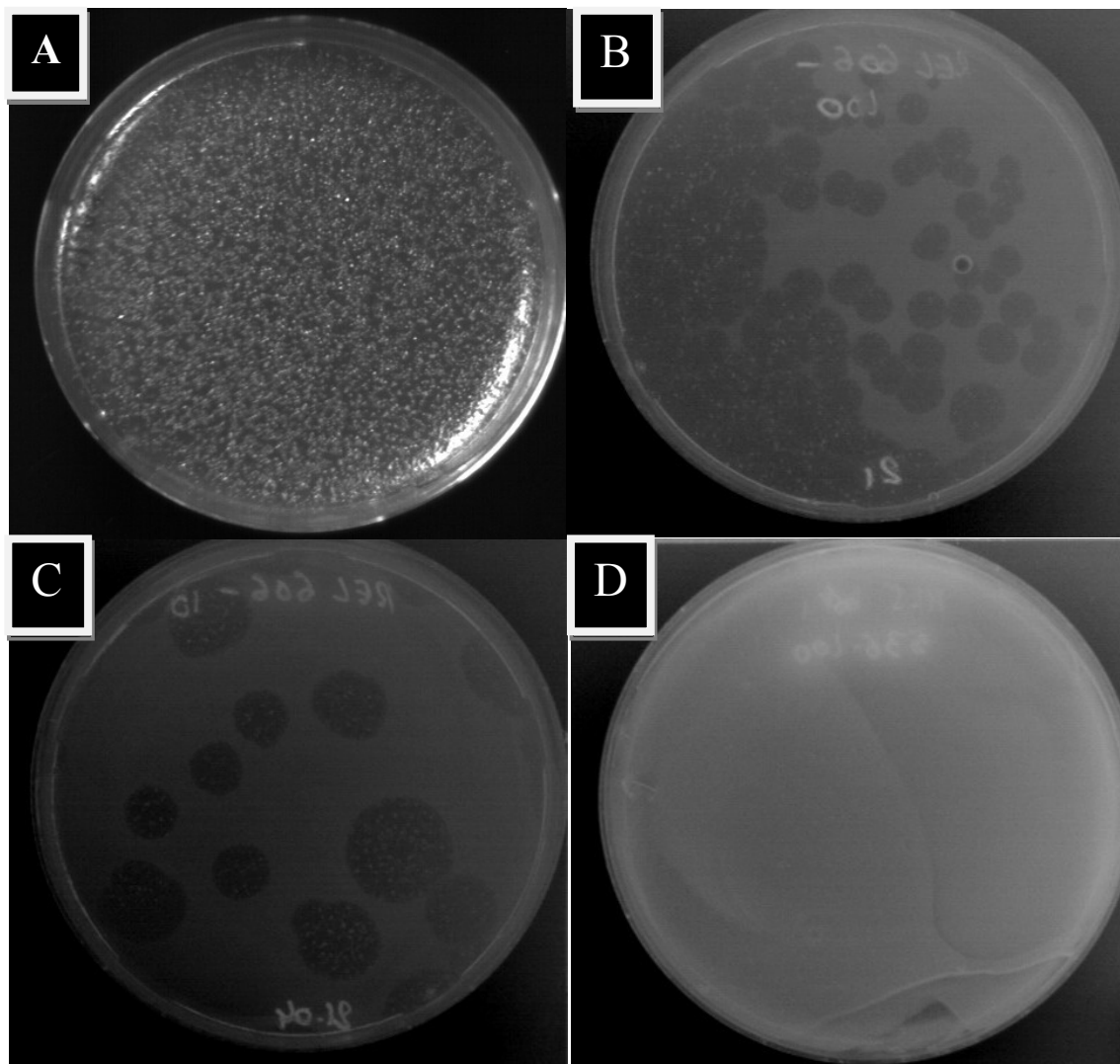


Figure 18: Observation des plages de lyse à J6. (A) souche avirulente b REL606 (10^8 UFC) incubées avec 10^3 amibes (B) avec 10^2 amibes (C) avec 10 amibes, phénotype dit Grazing sensible (GS). (D) souche virulente 536 (10^8 UFC) incubée avec 10^3 amibes, aucune plage de lyse n'apparaît, phénotype dit Grazing résistant (GR).

f). Protocole de pathologie expérimentale avec la souris :

Des bactéries en phase de croissance exponentielle ayant cultivé en bouillon LB pendant 16h sous agitation à 37°C sont lavées deux fois dans 0,9% de NaCl. Les bactéries sont remises en solution dans du sérum physiologique avant injection à la concentration de 10^9 UFC/ml . Les souris reçoivent une injection sous-cutanée de 0,2 ml de suspension bactérienne dans le cou.

Elles sont surveillées toutes les heures pendant 24h, puis pluriquotidiennement pendant une semaine. L'heure de leur mort est enregistrée. Les souris ayant survécu plus de 7 jours sont considérées comme guéries et sacrifiées (Picard *et al.*, 1999).

Les souris sacrifiées ou mortes sont congelées à -20°C pour étude ultérieure. Toutes les expériences sont répétées deux fois. Pour chaque souche, 10 souris sont inoculées. Pour mettre en évidence des différences statistiquement significatives, des effectifs plus grands de souris ont dû être testés (maximum de 20 souris supplémentaires).

Toutes les expérimentations animales ont été conduites dans le respect de la réglementation en vigueur.

Décompte bactériens au niveau de la rate:

Les rates sont ensuite extraites et pesées. Un millilitre de sérum physiologique est ajouté à chaque rate. Elles sont ensuite broyées et homogénéisées grâce à un broyeur de tissu (Ultra-Turrax 25, Fisher Scientific®, Strasbourg, France).

Les comptages bactériens sont effectués sur géloses LB avec ou sans antibiotiques afin de distinguer les souches utilisées. Les résultats sont exprimés en Log_{10} d'UFC par mg de tissu.

g). Mesures du taux d'invasion intracellulaire :

Les co-cultures amibes-bactéries sont incubées avec un antibiotique, la gentamycine ($5\text{ }\mu\text{g/ml}$) pendant 45 mn, 2h et 3h.

Les bactéries mortes et l'antibiotique en suspension dans le milieu sont éliminés par centrifugation à 2000 tr/mn pendant 15 mn à 4°C . Le culot d'amibes est resuspendu dans 5ml de MCPB. On estime la densité de la population amibienne par décompte sur Kovaslide®.

Les membranes des amibes sont rompues par l'ajout de $2,5\text{ }\mu\text{L}$ d'un détergent concentré (Sodium Dodécyl Sulfate 20%) pendant une durée de 2 minutes dans 1 ml de culture.

Les bactéries intracellulaires sont ensuite récupérées dans 1 ml de MCPB après rinçage par centrifugation de 15 min à 10 000 rpm à 4°C. Un volume de 100 µl de cette suspension est utilisé pour réaliser des dilutions de 10^{-1} à 10^{-6} . Ces dilutions sont étalées sur boîte de Pétri contenant de la gélose LB. Le taux de bactéries intracellulaires est estimé à partir du décompte des colonies sur ces boîtes après 24 h d'incubation à 37°C.

h). Analyses statistiques :

Les analyses statistiques, les courbes de Kaplan-Meier ont été effectuées grâce au logiciel XLSTAT®.

III. Résultats

L'objectif de ce travail est de préciser l'impact du fond génétique des souches de *E. coli* sur la virulence liée au HPI. Dans ce but, trois souches de *E. coli* 536, CFT073 et NU14 appartenant toutes au groupe phylogénétique B2, et ont été choisies comme représentatives de trois sous-groupes différents, les sous-groupes II, III et IX qui regroupent majoritairement les souches d'ExPEC. Ces trois souches ont été utilisées dans toutes les expériences. Par la suite et au vu des premiers résultats chez la souche NU14, deux autres souches du sous-groupe IX de B2 ou de ST 95 selon le schéma d'Achtman (souches de méningites néonatales C5 et S88) ont été testées. Ce sont les souches génétiquement les plus proches de NU14, en particulier la souche C5 qui appartient au même complexe clonal O18:K1:H7.

1). Détermination des facteurs de virulence :

L'étude des FVs permet d'évaluer la virulence intrinsèque de chacune des souches. Un score de virulence a été calculé correspondant à la somme des FV (adapté de Lefort et coll. 2011).

La séquence complète du génome de NU14 n'étant pas encore disponible, la présence de gènes de virulence caractéristiques des ExPEC a été recherchée par PCR et comprend : des adhésines (*papC*, *papG*, incluant les allèles de *papG* qui sont *papGII* et *papGIII* ; *sfa/foc*, *hra*, *iha* et *ibeA*), des toxines (*hlyC*, *cnf1*, *sat*, *usp* et *clbQ*), des systèmes de captation du fer (*fyuA*, *irp2*, *iroN*, *iucC* et *ireA*), des protectines (*neuC*, *ompT* et *TraT*), ainsi qu'un gène codant pour une protéine uropathogène spécifique *usp*. Les souches 536, CFT073 et S88 étant entièrement séquencées, la présence de ces gènes était recherchée par BLAST sur NCBI avec vérification de l'intégrité du gène (<http://www.ncbi.nlm.nih.gov>) (Tableau XII). Les souches 536 et S88 avaient un score de virulence de 12 et 13 *versus* 15 et 16 pour CFT073 et NU14. Chez 536, il

n'y avait pas de gènes associés aux méningites néonatales (tels que *neuC* codant pour la capsule K1 et *ibeA*). Quant à la souche S88, elle apparaît relativement pauvre en FV.

Tableau XII : Principaux FV portés par les souches de *E. coli* 536, NU14, CFT073 et S88.

Gènes de virulence											
Souches	<i>chuA</i>	<i>sfa/foc</i>	<i>iroN</i>	<i>aer</i>	<i>papC</i>	<i>papG</i>	<i>hlyC</i>	<i>cnf1</i>	<i>hra</i>	<i>fyuA</i>	<i>irp2</i>
536	+	+	+	-	+	III	+	-	+	+	+
NU14	+	+	+	-	+	III	+	+	+	+	+
CFT073	+	+	+	+	+	II	+	-	-	+	+
S88	+	-	+	+	+	II	-	-	-	+	+

Gènes de virulence											
Souches	<i>kpsE</i>	<i>ireA</i>	<i>ibeA</i>	<i>usp</i>	<i>traT</i>	<i>neuC</i>	<i>iha</i>	<i>sat</i>	<i>ompT</i>	<i>clbQ</i>	Score*
536	+	-	-	+	-	-	-	-	+	+	12
NU14	+	-	+	+	+	+	-	-	+	+	16
CFT073	+	+	-	+	-	-	+	+	+	+	15
S88	+	+	-	+	+	+	-	-	+	+	13

* : score de virulence, calculé comme la somme des FV, chacun affecté à un point.

iroN : salmochéline, *aer* : aérobactine, *fyuA-irp2* : yersiniabactine.

2). Virulence chez la souris dans un modèle de septicémie :

Les souches ont été testées dans un modèle murin de septicémie (Picard 1999). Ce modèle complexe de virulence étudie le passage des bactéries dans le sang et leur dissémination secondaire, et ce avec l'entrée en action de plusieurs lignes de défense immune chez l'hôte infecté.

a). Comparaison des souches sauvages :

Dans un premier temps, les courbes de survie des souris inoculées par les souches sauvages 536, CFT073, NU14, C5 et S88 ont été comparées par le test du Log-Rank (Tableau XIII).

Tableau XIII : Valeurs des tests de Log-Rank (*p*) entre les courbes de survie des amibes inoculées par les souches 536, CFT073, S88, NU14 et C5. En gras sont indiqués les *p* significatifs.

Souches	C5	S88	NU14	CFT073
536	<0.001	0.003	<0.001	<0.0001
CFT073	0.873	0.007	<0.0001	
NU14	0.001	<0.0001		
S88	0.001			

Les souches par ordre de virulence décroissante sont : la souche 536 qui tue toutes les souris en moins de 20h (dont 95% en moins de 16h), suivie par la souche S88 qui tue 100% des souris en moins de 21h, puis CFT073 et C5 (de survies non différentes statistiquement) qui tuent 100 % des souris en environ 40 h, et enfin la souche NU14 avec 90% de mortalité en 5 jours (Figure 19). On constate que la virulence de ces souches dans le modèle de septicémie est sensiblement différente d'une souche à l'autre (Figure 19).

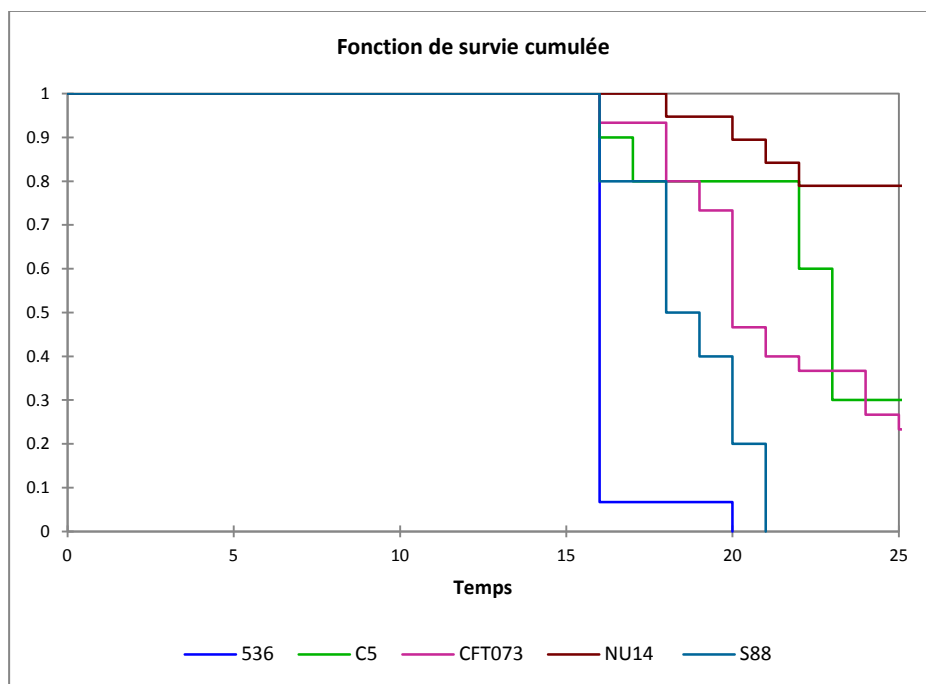


Figure 19 : Courbes de Kaplan-Meier des souris inoculées par les souches 536, C5, CFT073, NU14 et S88. En ordonnées sont indiqués les pourcentages de morts, en abscisses les heures de mort des souris.

b). Souche 536 et ses mutants du HPI :

La souche 536 Δ HPI avait déjà été testée dans l'étude de Médéric Diard (Diard 2010) montrant une absence d'atténuation de la virulence par rapport à la souche sauvage. Devant ce résultat, les mutants des trois gènes du HPI de cette souche n'ont pas été testés chez la souris.

c). Souche CFT073 et ses mutants du HPI :

La souche de *E. coli* CFT073 est un « mutant naturel » d'*irp2* par insertion d'une IS. Elle n'utilise pas la yersiniabactine comme sidérophore, mais possède un HPI complet avec YbtA comme possible régulateur pleiotrope.

La souche CFT073 Δ HPI::Cm inoculée aux souris avait une courbe de survie non statistiquement différente de celle de la souche sauvage (test du Log Rank $p=0.276$) (Figure 20), de même que le mutant CFT073 *ybtA*::Kn (donnée non montrée).

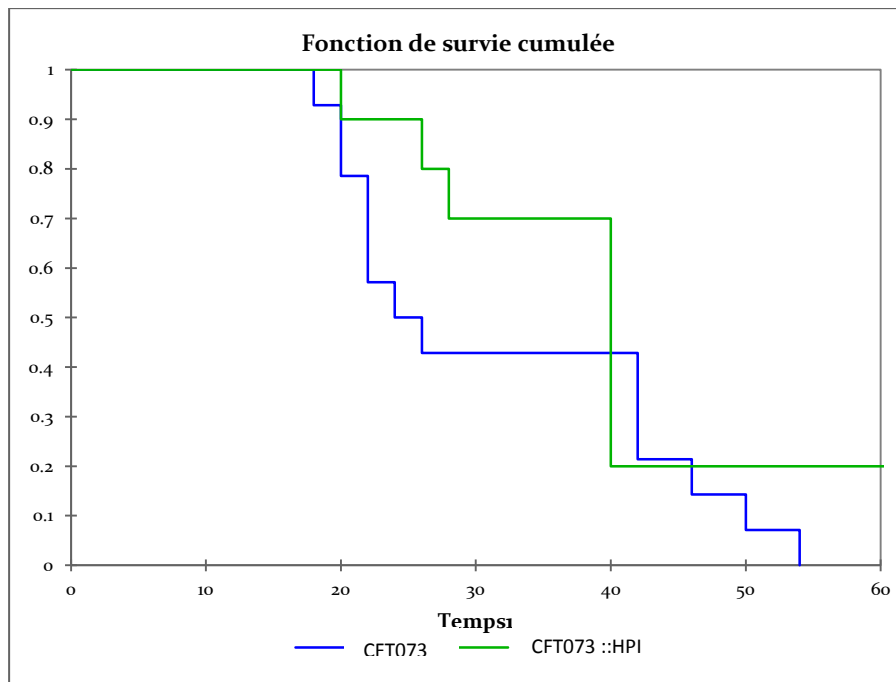


Figure 20 : Courbes de Kaplan-Meier des souris inoculées par la souche CFT073 et par son mutant du HPI (CFT073ΔHPI::Cm) dans un modèle de septicémie. En ordonnées sont indiqués les pourcentages de morts, en abscisses les heures de mort des souris (test du Log Rank $p=0.276$).

d). Souche NU14 et ses mutants du HPI :

La souche NU14 est un « tueur lent » avec des morts de souris étalés sur 5 jours et une mortalité de 90 %.

Le mutant de NU14 délété du HPI complet (NU14 HPI::Kn) présentait une baisse drastique de virulence (10% de mortalité *versus* 90% pour la souche sauvage) ($p<0.00001$) (Figure 21).

Les deux mutants de NU14 du gène de structure de la yersiniabactine *irp2* (NU14 *irp2*::Cm et NU14 *irp2*::Kn) construits lors de deux expériences indépendantes de mutagenèse par la méthode de Datsenko-Wanner étaient dénués de virulence et n'entraînaient aucune mort chez la souris (0% de mortalité) (Figure 22). Toutefois, il n'y avait pas de différence

statistiquement significative entre les mutants de *irp2* et celui du HPI (test du Log-Rank =0.157).

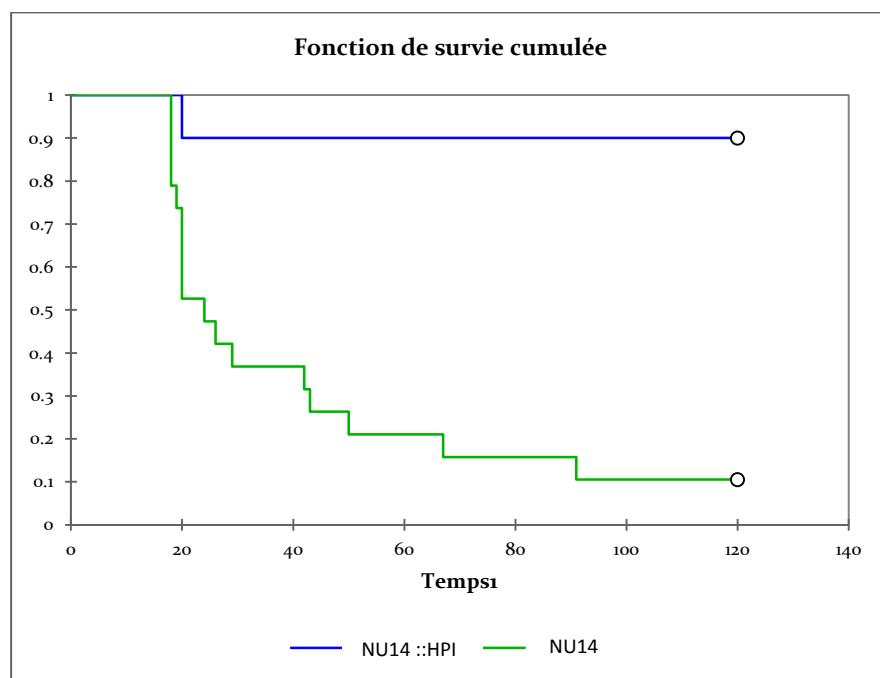


Figure 21 : Courbes de Kaplan-Meier des souris inoculées par la souche NU14 et par son mutant du HPI (NU14 HPI::Kn) dans un modèle de septicémie. En ordonnées sont indiqués les pourcentages de mort, en abscisses les heures de mort des souris (test du Log rank < 0.0001).

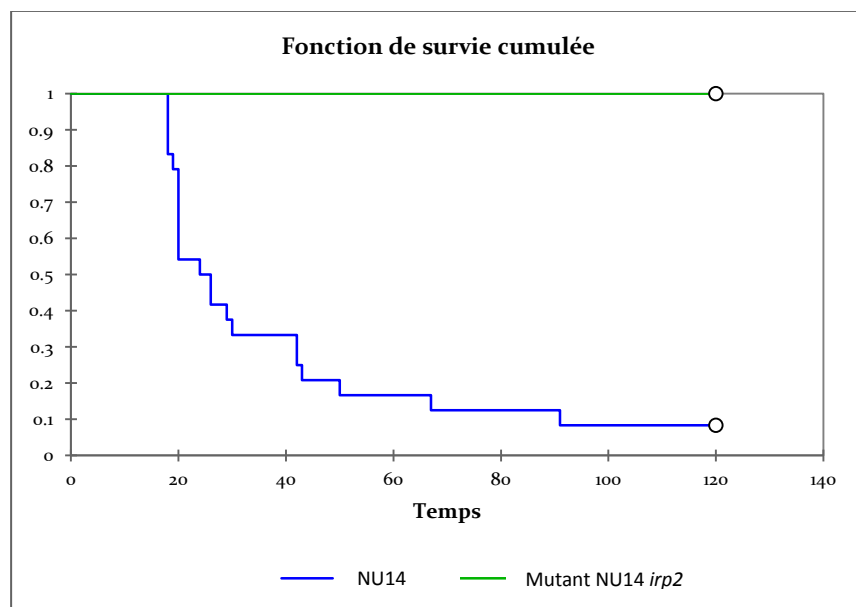


Figure 22 : Courbes de Kaplan-Meier des souris inoculées par la souche NU14 et par ses mutants du gène *irp2* (NU14 *irp2*::Cm et NU14 *irp2*::Kn) dans un modèle de septicémie. En ordonnées sont indiqués les pourcentages de décès, en abscisses les heures de décès des souris ($p<0.0000$).

Les mutants de NU14 de *ybtA* (gène du régulateur) construits par la méthode de Datsenko-Wanner lors de deux expériences différentes (NU14 *ybtA*::Kn 1 et NU14 *ybtA*::Kn 2) étaient moins virulents que la souche sauvage avec seulement 60 % de mortalité (test du Log-Rank, $p=0.015$) (Figure 23). L'effectif de souris nécessaire à cette expérience a été de 20 souris au lieu de 10 pour le test standard.

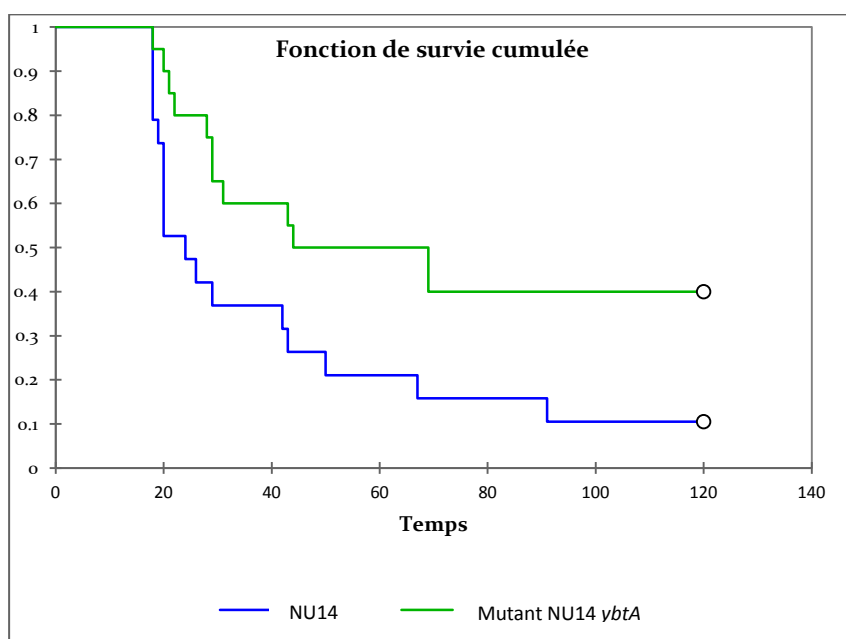


Figure 23 : Courbes de Kaplan-Meier des souris inoculées par la souche NU14 et par ses mutants du gène *ybtA* (NU14 *ybtA*::Kn 1 et NU14 *ybtA*::Kn 2) dans un modèle de septicémie.

En ordonnées sont indiqués les pourcentages de mort, en abscisses les heures de mort des souris.

Le mutant de NU14 de *fyuA* (gène du récepteur) (NU14 *fyuA*::Cm) tuait environ 50% des souris, et était significativement moins virulent que la souche sauvage (test du Log-Rank, $p<0.001$) (Figure 24). L'effectif de souris nécessaire à cette expérience a également été de 20

souris au lieu de 10. Ce phénotype de virulence atténuée se rapproche de celui du mutant de *ybtA* de NU14.

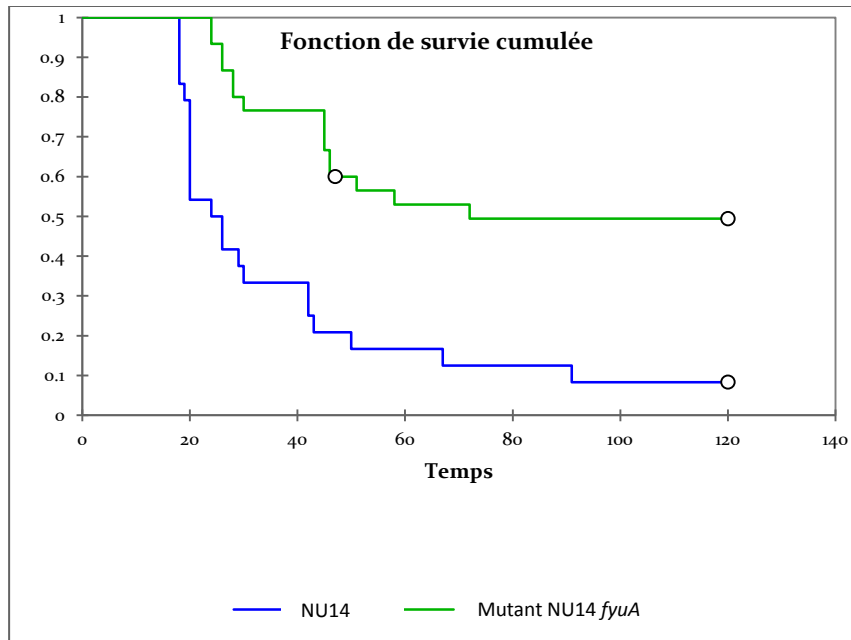


Figure 24 : Courbes de Kaplan-Meier des souris inoculées par la souche NU14 et par son mutant du gène *fyuA* (NU14 *fyuA*::Cm) dans un modèle de septicémie.

En ordonnées sont indiqués les pourcentages de mort, en abscisses les heures de mort des souris ($p < 0.001$).

Le mutant de NU14 de *ybtA* recomplémenté par le plasmide pWSK30 (NU14 *ybtA*::Kn pWSK30-*ybtA*) avait la même virulence que la souche mutée (donnée non montrée). La recomplémentation phénotypique avec restauration de la virulence n'a pas été possible. Ceci a également été le cas pour les recomplémentations plasmidiques des mutants de *fyuA* et d'*irp2* (NU14Δ*irp2* pWKS30-*irp2* et NU14 *fyuA*::Cm pACYC184-*fyuA* respectivement) ainsi qu'avec la souche portant un vecteur d'intégration chromosomique pCP1 (NU14 *ybtA*::*kn* pCP1). Les rates des souris mortes ont été étalées à la recherche de bactéries et ont montré que celles-ci avaient perdu le marqueur de résistance du vecteur pCP1 suggérant un

évènement de recombinaison *in vivo* peut-être induit par le stress auquel ont été soumises ces bactéries.

e). Souche C5 et ses mutants du HPI:

Après avoir comparé les trois souches virulentes appartenant à trois différents sous-groupes de B2 NU14, CFT073 et 536, et constatant que l'effet du HPI sur la virulence ne se manifeste que chez NU14, nous avons été amenés à étudier l'impact du HPI sur la virulence des souches C5 et S88 appartenant au sous-groupe IX comme NU14, et pour C5 appartenant au même complexe clonal 018:K1:H7.

La souche C5 délétée du HPI (C5Δ HPI::*cat-I*) avait une virulence significativement atténuée chez la souris par rapport à la souche sauvage (Figure 25).

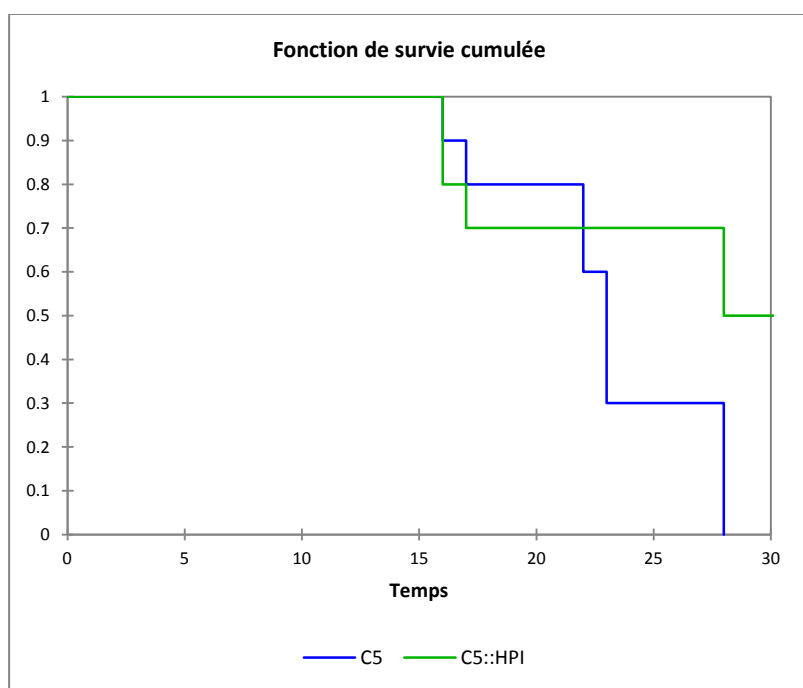


Figure 25 : Courbes de Kaplan-Meier des souris inoculées par la souche C5 et par son mutant du HPI (C5Δ HPI::*cat-I*) dans un modèle de septicémie. En ordonnées sont indiqués les pourcentages de mort, en abscisses les heures de mort des souris (test du Log Rank $p=0.037$).

Par contre, l'inoculation de C5 *fyuA::cat-1* n'entraînait pas de modification de mortalité par rapport à la souche sauvage (test du Log rank, $p=0.278$) (Figure 26). Le test standard comportant seulement 10 souris, un effectif plus grand pourrait peut-être mettre en évidence une différence significative, ainsi pour les mutants de NU14 de *fyuA* ou de *ybtA*, des effectifs de 20 souris avaient été nécessaires.

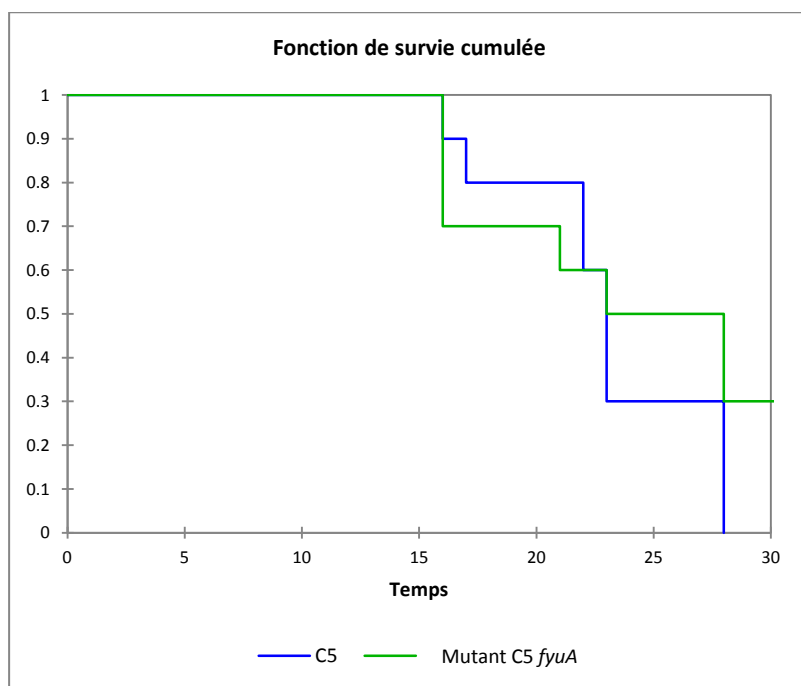


Figure 26 : Courbes de Kaplan-Meier des souris inoculées par la souche C5 et par son mutant de *fyuA* (C5Δ *fyuA::cat-1*) dans un modèle de septicémie. En ordonnées sont indiqués les pourcentages de mort, en abscisses les heures de mort des souris (test du Log Rank $p=0.278$).

f). Souche S88 et ses mutants du HPI

Les souris inoculées par le mutant du HPI de la souche S88 (S88 ΔHPI::*cat-1*) avaient une survie non différente de celles inoculées par la souche sauvage. Toutefois le test du Log rank avec un $p=0.09$ suggère que des effectifs plus grands que les 10 souris qui ont été testées pourrait mettre en évidence une différence (Figure 27). Il n'y avait pas de différence de survie entre S88 et son mutant *fyuA* (S88 *fyuA::cat-1*).

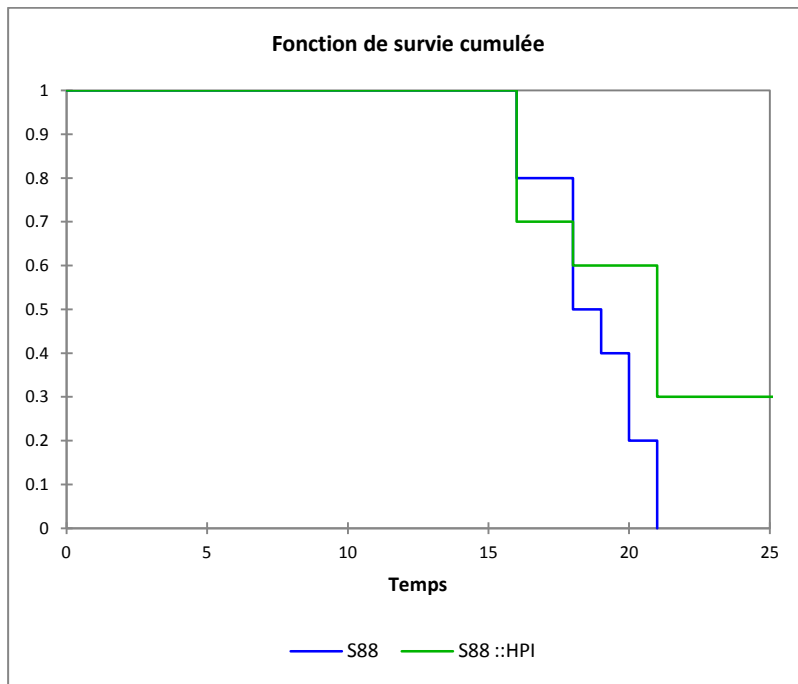


Figure 27 : Courbes de Kaplan-Meier des souris inoculées par la souche S88 et par son mutant du HPI (S88 Δ HPI::cat-1) dans un modèle de septicémie. En ordonnées sont indiqués les pourcentages de mort, en abscisses les heures de mort des souris (test du Log Rank $p=0.09$).

Ci-dessous, un tableau récapitulatif des effets chez la souris des différentes mutations :

	HPI	<i>irp2</i>	<i>ybtA</i>	<i>fyuA</i>
536	-	ND	ND	ND
CFT073	-	Mutant naturel	-	ND
NU14	++	++	+	+
C5	+	ND	ND	-
S88	+/-	ND	ND	-

ND : non déterminé.

3). Résultats du modèle de virulence expérimentale chez l'amibe :

Le modèle amibe est peu cher et reproductible et éthiquement sans limite par rapport aux souris. Il nous a ainsi permis de tester toutes les souches du tableau X avec au moins deux expériences indépendantes et au minimum trois réplicats par souche et par expérience. Les souches NU14, CFT073, 536, S88 et C5 étaient toutes virulentes chez l'amibe avec des phénotypes GR (résistants au broutage) à la densité de 10^3 , 10^2 , 10^1 amibes pour 10^8 UFC de bactéries (Figure 18, D).

a). Mutants du HPI de la souche de *E. coli* 536 :

La souche 536 délétée du HPI complet (536 Δ HPI) présentait le même phénotype GR que la souche sauvage et était donc virulente chez l'amibe. De même, les mutants de 536 des gènes *fyuA*, *ybtA* ou d'*irp2* (soit 536 *fyuA*::Cm, 536 *ybtA*::Kn et 536 *irp2*::Kn, respectivement) présentaient le phénotype GR.

Le mutant de 536 délété des PAIs III et IV (536 Δ PAI III-HPI) (soit ne sécrétant ni la yersiniabactine du HPI, ni la salmochéline portée par le PAI III et ayant ainsi pour unique sidérophore majeur l'entérobactine) restait hautement virulent chez l'amibe de même que la souche 536 délétée de ses sept PAIs.

Il est à noter que la souche 536 possède le même arsenal de sidérophores que la souche NU14.

b). Mutants du HPI de *E. coli* CFT073 :

La souche de *E. coli* CFT073 n'utilise pas la yersiniabactine comme sidérophore, mais possède un HPI avec YbtA comme possible régulateur pleiotrope. La délétion du HPI complet (CFT073 HPI::Kn) ou du gène codant pour YbtA (CFT073 *ybtA*::Kn) n'affectait pas sa virulence chez l'amibe.

c). Mutants du HPI de NU14 :

La délétion complète du HPI chez NU14 (NU14 HPI::Kn) le rendait sensible à la digestion par l'amibe dès 10^2 amibes pour 10^8 UFC de bactéries avec un phénotype GS correspondant à l'apparition de plages de lyse bactérienne (Tableau XIV).

D'autre part, les deux mutants indépendants de NU14 du gène de structure *irp2* NU14 *irp2::kn* et NU14 *irp2::Cm* présentaient des plages de lyse dès une densité d'amibes de 10 pour 10^8 UFC de bactéries, comme la souche avirulente B REL 606. Le phénotype virulent GR de la souche sauvage était restauré par recombplémentation plasmidique (NU14 Δ *irp2* pWKS30-*irp2*) (Tableau XIV).

Tableau XIV : Rôle de Irp2 et du HPI dans la résistance de NU14 à la digestion par l'amibe.

Souche	NU14	-----	-----	-----
	NU14 HPI::Kn	-----	++++--	++++++
	NU14 <i>irp2::Kn</i>	+++ - - -	+++ - - -	++++++
	NU14 Δ <i>irp2</i> pWKS30- <i>irp2</i>	-----	-----	-----+
	B REL 606	++++++	++++++	++++++
		10	10^2	10^3
		Densités d'amibes		

(+) représente le nombre réplcats avec plaques de lyses à 6 j après co-culture. (-) représente le nombre de réplcats sans plages de lyse.

Les deux mutants indépendants du régulateur YbtA (NU14 *ybtA::kn* 1 et 2) présentaient une apparition de plages de lyse en compétition avec une forte densité d'amibes à 10^3 (Tableau XV). Ce phénotype d'atténuation de virulence GS a été confirmé en testant les deux mutants construit de façon indépendante. De plus, la recombplémentation chromosomique par le vecteur suicide pCP1 a restauré le phénotype sauvage GR (NU14 *ybtA::kn* pCP1) (Tableau XV, Figure 28). La recombplémentation plasmidique avec le plasmide pWKS30-*ybtA* par contre n'a pas été possible et ce en maintenant ou non la pression de sélection antibiotique dans le modèle (données non montrées).

Tableau XV : Rôle de YbtA dans la résistance de NU14 à la digestion par l'amibe.

Souche	NU14	-----	-----	-----
	NU14 <i>ybtA::kn</i> (1 et 2)	-----	-----	++++++
	NU14 <i>ybtA::kn</i> pCP1	-----	-----	-----+
	REL 606	++++++	++++++	++++++
		10	10^2	10^3
		Densités d'amibes		

(+) représente le nombre réplcats avec plaques de lyses à 6 j après co-culture. (-) représente le nombre de réplcats sans plages de lyse.

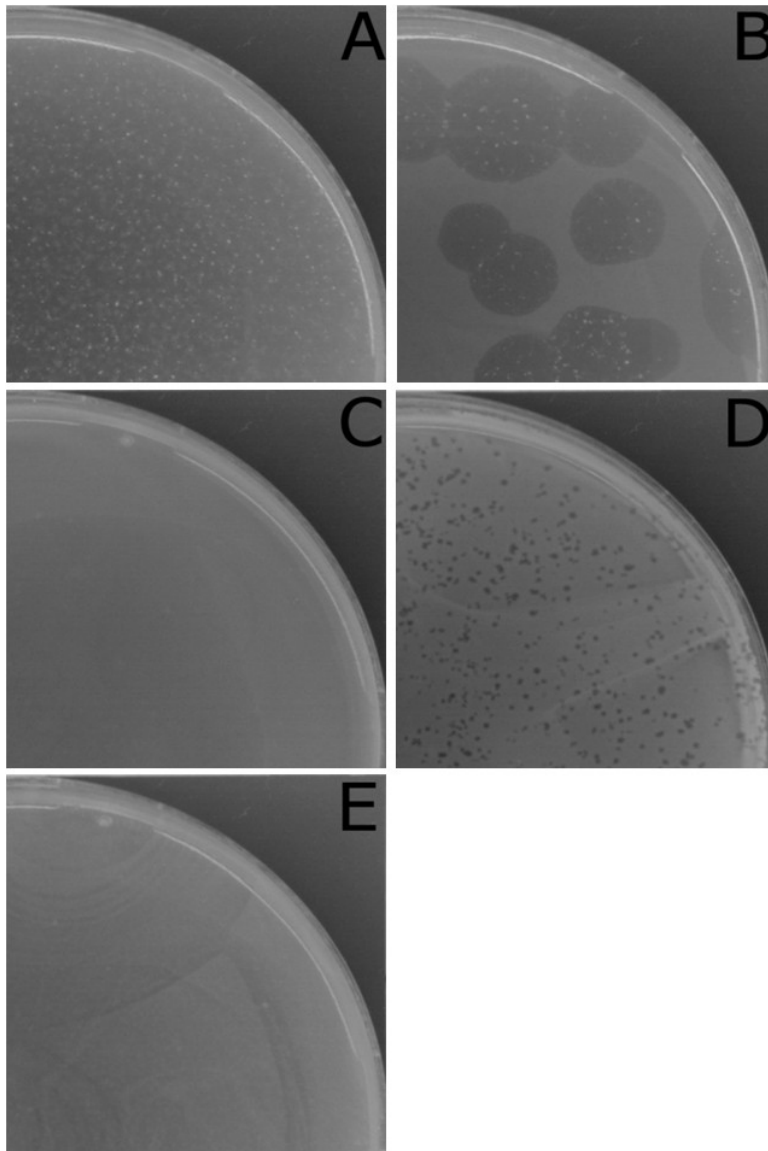


Figure 28 : Plages de lyse observées à J6.

- (A) *E. coli* B REL 606 (10^8) avec *D. discoideum* (10^3)
- (B) REL 606 avec *D. discoideum* (10^2)
- (C) NU14 (10^8) avec *D. discoideum* (10^3)
- (D) mutant NU14 *ybtA::kn* avec *D. discoideum* (10^3)
- (E) NU14 $\Delta ybtA$ pCP1 avec *D. discoideum* (10^3)

Le mutant de la souche NU14 du récepteur FyuA (NU14 *fyuA::Kn*) présentait une apparition de plages de lyse à 10^3 densité d'amibes et avait donc le même phénotype de virulence atténuée que le mutant du régulateur YbtA. La recomplémentation plasmidique a été

obtenue avec le plasmide à faible nombre de copies pACYC184-*fyuA* en additionnant de la tétracycline aux géloses HL5 (Tableau XVI).

Tableau XVI : Rôle de FyuA dans la résistance de NU14 à la digestion par l'amibe.

Souche	NU14	-----	-----	-----
	NU14 <i>fyuA</i> ::Kn	-----	-----	++++++
	NU14 <i>fuyA</i> ::Kn pACYC184- <i>fyuA</i>	-----	-----	-----
	REL 606	++++++	++++++	++++++
		10	10 ²	10 ³
		Densités d'amibes		

(+) représente le nombre réplicats avec plaques de lyses à 6 j après co-culture. (-) représente le nombre de réplicats sans plages de lyse.

d). Mutants du HPI de *E. coli* C5 :

Les mutants de la souche C5 du HPI et de *fyuA* (C5ΔHPI::*cat-I*, C5 *fyuA*::*cat-I*) étaient de moindre virulence avec apparition de plages de lyse à 10³ densité d'amibes. Par introduction du plasmide pACYC184-*fyuA*, le phénotype virulent a été restauré sans qu'il y ait eu nécessité de maintenir la pression de sélection par adjonction d'antibiotiques au milieu (tableau XVI). La souche C5 appartient au même sous-groupe de B2 que NU14, et au même complexe clonal O18:K1:H7, elle produit les mêmes sidérophores.

Tableau XVI : Rôle de FyuA et du HPI dans la résistance de la souche C5 à la digestion par l'amibe.

Souche	C5	-----	-----	-----
	C5 Δ HPI:: <i>cat-1</i>	-----	-----	--++++
	C5 <i>fyuA</i> :: <i>cat-1</i>	-----	-----	--++++
	C5 <i>fyuA</i> :: <i>cat-1</i> pACYC184- <i>fyuA</i>	-----	-----	-----
	REL 606	++++++	++++++	++++++
		10	10 ²	10 ³
		Densités d'amibes		

(+) représente le nombre réplicats avec plaques de lyses à 6 j après co-culture.

(-) représente le nombre de réplicats sans plages de lyse.

e). Mutants du HPI de *E. coli* S88 :

Les mutants de S88 délété du HPI complet (S88 Δ HPI::*cat-1*) ou muté du gène *fyuA* (S88 *fyuA*::*cat-1*) présentaient le phénotype GR et n'avaient pas d'atténuation de virulence chez l'amibe.

Cette souche appartient comme NU14 et C5 au sous-groupe IX de B2, mais est de complexe clonal différent (O45:K1:H7). Elle porte le plasmide pS88 codant pour la salmochéline et l'aérobactine, elle produit ainsi un sidérophore supplémentaire par rapport à NU14 et C5. Il reste à tester les mutants du HPI de S88 curés de ce plasmide afin de voir si la yersiniabactine devient alors critique pour la virulence de la souche dans ce modèle.

Ci-dessous, un tableau récapitulatif des effets chez l'amibe des différentes mutations :

	HPI	<i>irp2</i>	<i>ybtA</i>	<i>fyuA</i>
536	-	-	-	-
CFT073	-	Mutant naturel	-	ND
NU14	++	++	+	+
C5	+	ND	ND	+
S88	-	ND	ND	-

ND : non déterminé.

4). Mesures du taux d'invasion intracellulaire (intra-amibes) :

Le test consiste en une co-culture amibes-bactéries en excès de bactéries avec mesure des bactéries intra-amibiennes en deux temps à 1 h et à 3 h. Les trois souches étudiées sont 536, CFT073 et NU14.

Pour la souche témoin avirulente B REL606, aucune bactérie intra-amibe n'est retrouvée à 1h et 3h, ce qui pourrait traduire une digestion très rapide des bactéries par l'amibe.

Les souches virulentes chez l'amibe NU14, CFT073 et 536 étaient retrouvées au cours du temps en intra-amibe mais avec des différences notables selon les souches. Les deux souches CFT073 et 536 sont de plus en plus internalisées au cours du temps (facteur 5 à 10 à 3h par rapport à 1h) alors que la souche NU14 est moins internalisée à 3h qu'à 1h, suggérant une virulence par résistance à la phagocytose ou par un mécanisme secondaire d'échappement de la bactérie par rapport à l'amibe (Figure 29). Ces expériences ont été reproduites deux fois.

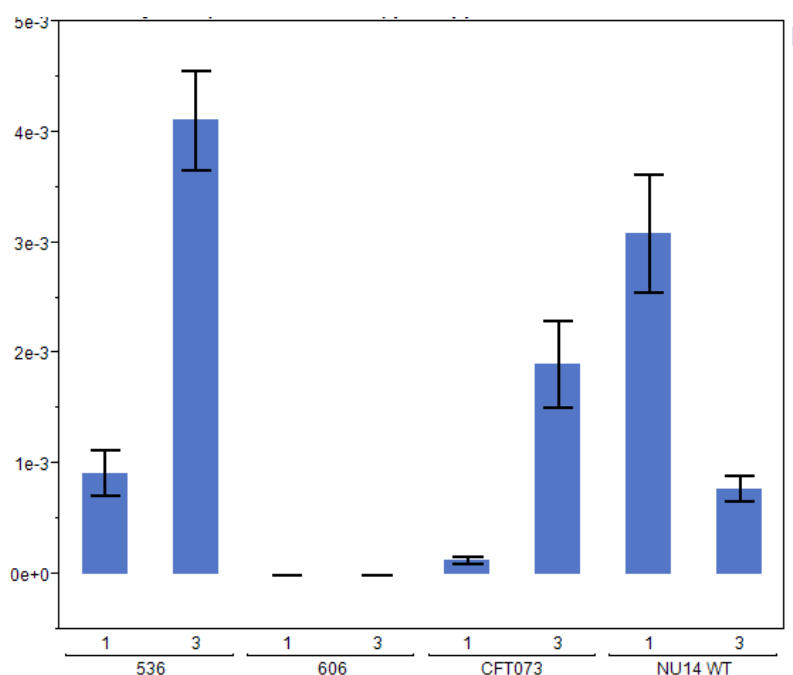


Figure 29 : Internalisation des bactéries dans les amibes à 1h et 3 h. En ordonnées sont indiquées les UFC internalisées par amibe, en abscisses les souches avec les deux temps de prélèvement.

Devant ce comportement différent de NU14, les différents mutants du HPI et des 3 gènes et leurs recombplémentés de la souche NU14 ont été testés dans ce modèle. La souche sauvage était donc moins internalisée au cours du temps, le mutant délété du HPI (NU14 Δ HPI::Kn) présentait la tendance inverse et était plus internalisé au cours du temps ($p=0,005$). Les mutants de NU14 FyuA, Irp2 et YbtA (NU14 *fyuA*::Kn, NU14 *irp2*::Kn, NU14 *ybtA*::Kn 1) présentaient également une tendance inverse avec augmentation de l'entrée intra-amibes entre 1h et 3h. De plus, cette tendance était recombplémentée dans tous les cas (souches NU14 *fyuA*::Kn pACYC184-*fyuA*, NU14 Δ *irp2* pWSK30-*irp2*, NU14 *ybtA*::Kn pCP1) (Figure 21).

A l'opposé, il n'y avait pas d'effet de la délétion du HPI sur l'internalisation de 536 ou de CFT073 ($p=0,10$ et $p=0,58$).

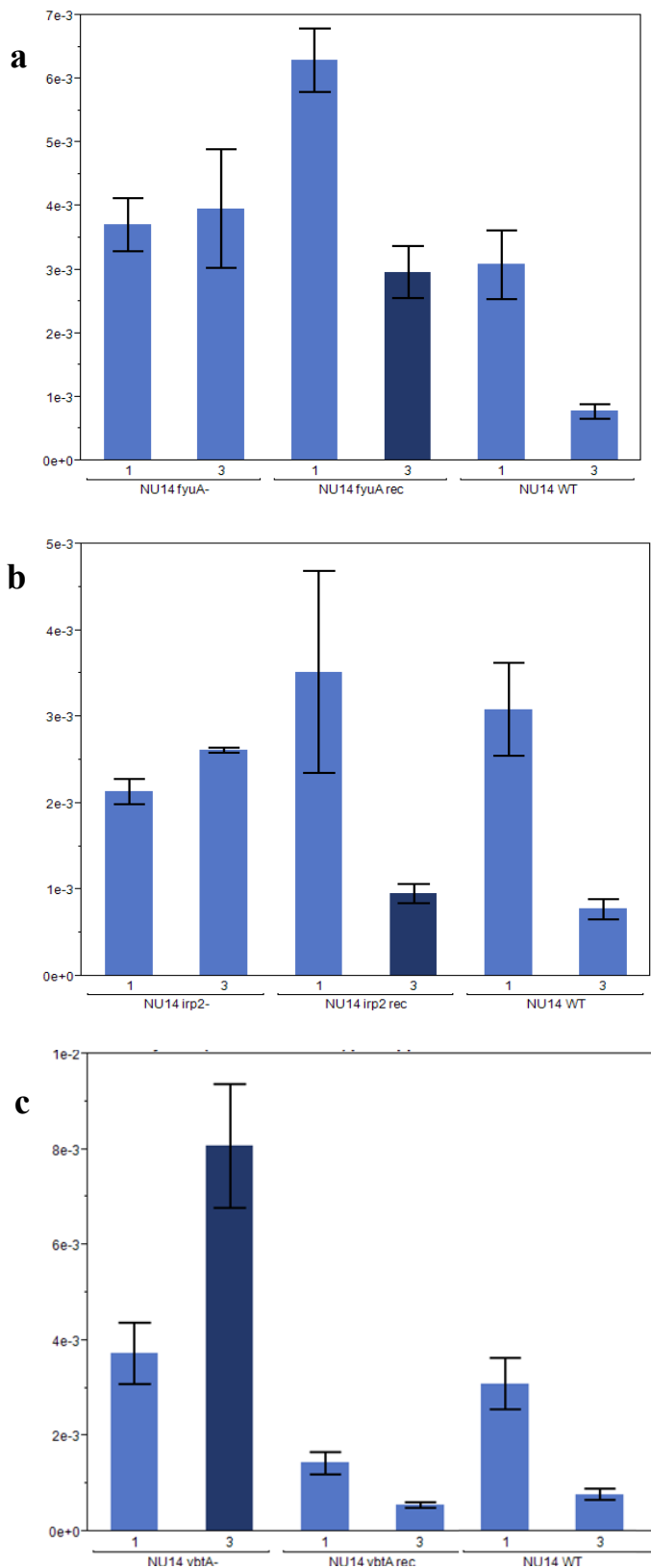


Figure 21 : Internalisation des bactéries dans les amibes à 1h et 3 h. En ordonnées sont indiquées les UFC internalisés par amibe, en abscisses les souches avec les deux temps de prélèvement. **(a)** : souches NU14, mutant NU14 *fyuA*::Kn (NU14 *fyuA*⁻) et recomplémenté

NU14 *fyuA*::Kn pACYC184-*fyuA* (NU14 *fyuA* rec), **(b)** souches NU14, mutant NU14 *irp2*::Kn (NU14 *irp2*-) et recomplémenté NU14 Δ *irp2* pWSK30-*irp2* (NU14 *irp2* rec), **(c)** souches NU14, mutant NU14 *ybtA*::Kn 1 (NU14 *ybtA*-) et recomplémenté NU14 *ybtA*::Kn pCP1 (NU14 *ybtA* rec).

Dans le but d'élucider le mécanisme de virulence de *E. coli* chez l'amibe, des expériences de co-culture d'amibes avec des surnageants de culture filtrés de 536, CFT073, NU14 ont été entreprises. Ces surnageants n'ont pas plus affecté la croissance de l'amibe sur 6 h que la souche avirulente B REL 606 (données non montrées), la virulence de ces souches n'est pas probablement pas due à des facteurs sécrétés dans le milieu tels que des toxines.

Ci-dessous, un tableau récapitulatif des effets sur l'internalisation dans l'amibe entre 1h et 3h des différentes mutations :

	sauvage	HPI	<i>irp2</i>	<i>ybtA</i>	<i>fyuA</i>
536	↑	↑	↑	↑	↑
CFT073	↑	↑	Mutant naturel	↑	ND
NU14	↓	↑*	↑*	↑*	↑*

ND : non déterminé. ↑ : augmentation du nombre de bactéries intra-amibes. ↓ : baisse du nombre de bactéries intra-amibes. * : différence statistiquement significative.

IV. Discussion:

Afin d'éclaircir le rôle du fond génétique dans l'expression de la virulence, trois isolats naturels appartenant au groupe phylogénétique B2 le plus représenté chez les ExPEC ont été choisis. Ces souches appartenaient à trois sous-groupes différents de B2 ayant une structure quasi-clonale avec très peu de STs différents par sous-groupe. Ces trois sous-groupes représentent à eux seuls plus de la moitié des souches de B2 (Le Gall *et al.*, 2007). Ce groupe phylogénétique est le plus pathogène chez l'homme, mais aussi comme nous l'avons trouvé dans notre précédente étude l'un des meilleurs colonisateurs de la flore intestinale. Nous avons choisi d'étudier l'impact de la perte d'un sidérophore, la yersiniabactine, facteur de virulence capital des bactéries dans leur rencontre avec l'hôte en milieu carencé en fer (Fetherston *et al.*, 2010) mais aussi facteur de commensalisme par captation du fer dans leur habitat naturel où elles sont en compétition avec d'autres micro-organismes (Perry *et al.*, 1999).

Dans notre travail, l'étude du contenu en FV des souches reflétant leur virulence intrinsèque apporte finalement peu de renseignements sur leur virulence *in vivo*. En effet, la souche NU14 bien qu'apparaissant mieux équipée (avec des FV de méningites néonatales tels que NeuC) que 536 est une souche «tueuse lente» chez la souris alors que 536 est une «tueuse rapide». La présence de FV ne présage pas de leur niveau de transcription, ni de leur expression *in vivo*. La seule acquisition de facteurs de virulence n'est probablement pas suffisante pour permettre l'émergence de la virulence, ce qui met en relief le rôle du fond génétique dans l'expression de celle-ci.

Cette virulence est mieux approchée par l'expérimentation *in vivo*. Le modèle de septicémie chez la souris est un modèle complexe et global qui reflète la confrontation entre les défenses immunes de la souris et la bactérie avec son matériel génétique de FVs, ses mécanismes de régulation et ses capacités d'adaptation métaboliques. Elle implique la translocation à partir

de l'inoculation sous-cutanée dans le sang, puis la multiplication de la bactérie durant la phase septicémique.

Dans le sang, la quantité de fer libre est contrôlée et maintenue très basse par fixation aux protéines plasmatiques comme décrit précédemment, c'est l'un des mécanismes de défense de l'hôte contre les bactéries (Miller *et al.*, 2006).

De cette première approche expérimentale, il ressort que la virulence des souches est différente au sein du même groupe phylogénétique B2 et y compris au sein du même sous-groupe IX avec par ordre décroissant les souches 536, S88, C5 et CFT073. NU14 apparaît comme la moins virulente.

Les deux souches des sous-groupes II et III soient CFT073 et 536 n'étaient pas affectées dans leur virulence par délétion du HPI ou d'un de ses gènes. CFT073 a la particularité de ne pas produire de yersiniabactine (par insertion d'IS) mais de porter un HPI complet, celui-ci n'a pas d'effet dans notre modèle de septicémie.

La souche 536 produit les sidérophores entérobactine, salmochéline et yersiniabactine mais pas l'aérobactine, elle n'est pas affectée par la délétion du HPI dans sa virulence.

A l'inverse, la souche NU14, qui produit les mêmes sidérophores majeurs que 536 voit sa virulence drastiquement diminuer lors de la délétion du HPI ou d'un de ses gènes. Les mutants du HPI complet ou du gène de synthèse *irp2* de NU14 étaient complètement avirulents chez la souris. Ceci pourrait traduire une moindre virulence de la souche NU14 qui serait plus facilement affectée par la perte d'un seul FV. Mais cela suggère aussi la notion de sidérotypage ou utilisation préférentielle d'un sidérophore en fonction de la situation de stress métabolique à laquelle est confrontée ces bactéries (ex : yersiniabactine et NU14). Cette hypothèse a été proposée dans l'étude de Garcia (Garcia *et al.*, 2010) où l'auteur indiquait que, pour une même souche, les sidérophores agiraient par niches métaboliques. Cette

utilisation préférentielle d'un sidérophore semble varier selon les souches dans notre travail (entre NU14 et 536).

D'autre part, dans le modèle souris, les mutants du gène de synthèse *irp2* étaient plus affectés dans leur virulence que les mutants du régulateur YbtA ou du récepteur FyuA.

Devant les résultats obtenus avec la souche NU14, d'autres souches du même sous-groupe IX, C5 et S88, et leurs mutants du HPI ont été testés. Le sous-groupe IX a en effet une structure quasi-clonale et ces souches ont le même ST, le ST95 selon le schéma d'Achtman. Ces souches se sont révélées plus virulentes que NU14 chez la souris, mais comme pour NU14, la délétion du HPI complet affectait leur virulence. Cette atténuation de virulence du mutant était différente entre C5 et S88. Elle est plus marquée pour C5, souche génétiquement la plus proche de NU14 et appartenant au même complexe clonal. Pour S88, cette atténuation est probable mais reste à confirmer. En effet, cette souche possède le plasmide pS88 qui code pour deux sidérophores, l'aérobactine et la salmochéline, et possède de ce fait un arsenal de sidérophores plus important que les souches NU14 et C5 (l'aérobactine). Ainsi, sur le même fond génétique un des événements d'acquisition par transfert horizontal de gènes de virulence et de gènes métaboliques est sans doute survenu. Chez S88, afin de préciser l'impact du fond génétique ou matériel génétique acquis verticalement *versus* le matériel génétique acquis horizontalement, il nous reste à tester dans le modèle de virulence murin différents mutants les souches S88 Δ pS88, S88 Δ HPI::*cat-1* Δ pS88, S88 *fyuA*::*cat-1* Δ pS88 et S88 *fyuA*::*cat-1* Δ pS88 pACYC184-*fyuA* correspondant à S88 et ses mutants du HPI curés de leur plasmide pS88 et recomplémentés afin d'étudier le rôle propre du fond génétique de la souche de *E. coli* S88 du sous-groupe IX dans la virulence liée au HPI.

La seconde partie étudie un modèle de virulence plus simple utilisant l'amibe qui est assimilable à un macrophage. L'amibe *D. discoideum* possède de nombreuses voies

communes avec celles des autres cellules eucaryotes. Elle permet d'étudier la phagocytose, l'une des lignes de défense des eucaryotes contre les bactéries.

Ce modèle moins onéreux et plus éthique que celui de la souris nous a permis de tester un grand nombre de mutants, et leurs recombplémentés. La recombplémentation a été possible dans tous les cas, le modèle permettant le maintien de la pression de sélection par adjonction d'antibiotiques au milieu. Celle-ci n'a pas toujours été nécessaire, le stress auquel sont soumis les bactéries est moindre dans ce modèle et une perte minime de fitness n'affecte pas globalement leur virulence sur 6 jours en co-culture avec des amibes. Le temps de division des amibes étant de 12h, les bactéries ont le temps de s'adapter à leur prédateur, à l'inverse de la souris où les premiers temps sont cruciaux.

Par contre, la recombplémentation n'a pas été possible chez la souris par perte du plasmide ou du vecteur inséré, ou par une modification même légère de fitness des souches lors de la mutagenèse, modification qui serait très délétère chez la souris. Cependant, des mutants construits indépendamment des gènes *irp2* et *ybtA* de NU14 ont donné les mêmes résultats. De même, la délétion des gènes du HPI ou du HPI allait dans le même sens d'atténuation de la virulence.

Une corrélation entre la virulence chez la souris et l'amibe a été mise en évidence comme décrit précédemment (Adiba *et al.*, 2010).

Les rôles des gènes du HPI ont pu être précisé dans ce modèle. De même que chez la souris, les mutants du HPI complet étaient quasi-avirulents. Les mutants du gène de synthèse *irp2* (ne produisant pas de yersiniabactine) se comportaient comme les mutants du HPI complet (tuant 0 souris sur 10, GS à une faible densité d'amibes), à l'inverse les mutants du régulateur YbtA ou du récepteur FyuA étaient de virulence légèrement atténuée seulement (tuant 5 à 6 souris sur 10, GS à une forte densité d'amibes).

Le mutant du récepteur FyuA synthétise la yersiniabactine mais ne peut l'internaliser complexée au fer et de ce fait ne peut l'utiliser comme sidérophore, bien qu'on ne puisse exclure l'hypothèse d'une entrée par une voie alterne inconnue pour le moment.

Le mutant du régulateur YbtA présente une baisse drastique de la production de yersiniabactine (Giuseppe Magistro, communication personnelle), celle-ci reste toutefois produite en très petites quantités.

Ces observations sont en faveur d'un rôle de la yersiniabactine en dehors de la captation du fer, de très petites quantités de celle-ci participant alors à la virulence de la souche. Elle pourrait agir seule ou complexée au fer en tant que molécule de signalisation et induire la transcription d'autres FVs situés en dehors du HPI sur le génome. Ou bien la yersiniabactine pourrait agir en ayant un effet de toxicité directe sur les macrophages par inhibition des ROS (Paauw *et al.*, 2009). Cette hypothèse est d'autant plus intéressante que l'amibe, organisme unicellulaire, est comme les macrophages un phagocyte professionnel. Toutefois, cette observation a été faite en utilisant de la yersiniabactine purifiée, et rien ne prouve que cela est transposable *in vivo*.

Pour expliquer les différences de virulence observées entre les souches 536, CFT073 et NU14 des expériences d'internalisation dans l'amibe ont été conduites et ces souches ont montré des comportements très différents.

La virulence de la souche NU14 serait liée à une moindre phagocytose par résistance à l'internalisation ou par échappement de l'amibe. Les souches de 536 et de CFT073 tueraient l'amibe par internalisation massive. Ce mécanisme de résistance à la phagocytose chez l'amibe a déjà été décrit pour *Klebsiella pneumoniae* via son lipopolysaccharide (March *et al.*, 2013). La souche NU14 appartient au groupe des souches de méningites néonatales et possède la capsule K1, qui pourrait être impliquée dans cette résistance à la phagocytose. Selon nos données, le HPI est indispensable à la résistance à la phagocytose de NU14. Il semblerait que

chez NU14, c'est la présence de ces deux facteurs (la capsule et le HPI) qui est nécessaire à sa virulence chez l'amibe. Cette observation pourrait être artéfactuelle par stimulation d'autres sidérophores ou d'autres facteurs de virulence chez ces mutants, avec des régulations complexes. Ces hypothèses restent à confirmer par microscopie confocale analysant l'entrée, la phagocytose, et l'éventuelle ressortie de la bactérie afin de préciser son mécanisme exact de virulence. Une collaboration avec l'équipe du Pr. Pierre Cosson de l'Université de Genève est prévue dans ce domaine.

CONCLUSION GENERALE :

Ces travaux ont permis la mise au point d'une méthode rapide, fiable et pratique de quantification de *E. coli* et de ses phylogroupes majeurs, permettant de traiter un grand nombre d'échantillons à la fois. Cette approche originale de PCR quantitative intra-espèce a confirmé les données de littérature sur les prévalences de *E. coli* chez l'homme et les animaux. L'homme est porteur dans près de 100% de *E. coli* dans son microbiote intestinal et à des quantités de près de 10^8 UFC/gr de fèces, plus élevées que chez les animaux sauvages et domestiques étudiés. De plus, nos études confortent l'hypothèse du clone dominant c'est à dire d'un clone représentant plus de 50% de la population de l'espèce que ce soit chez l'homme ou bien l'animal.

La méthode moléculaire mise au point a un seuil de détection très bas, et a ainsi révélé la présence de groupes minoritaires non détectables par les techniques phénotypiques précédentes. Ainsi, plus de 40% des sujets de l'étude Coliville hébergeaient ces clones minoritaires. Cette partie immergée de l'iceberg reste peu étudiée et expliquerait les différences de portage observées au cours du temps ou l'acquisition apparente de nouveaux clones qui seraient en fait des clones minoritaires qui émergeraient à l'occasion de pressions de sélection sur le microbiote. Cette technique donne aussi une estimation de la diversité de la population commensale de *E. coli* via l'étude de ses groupes phylogénétiques.

La suite du travail a consisté à collecter des selles d'animaux sauvages et domestiques issus de la même région géographique, l'Île-de-France. L'étude comparative des structures des populations commensales des hommes et des animaux sauvages et domestiques a révélé la présence d'entérocotypes, correspondant à des associations préférentielles de phylogroupes au sein des fèces. Comme décrit précédemment à l'échelle du microbiote intestinal total (Arumugam *et al.*, 2011), l'espèce de l'hôte, son régime et son habitat ont un impact important sur les entérocotypes à l'échelle de l'espèce bactérienne. Trois principaux

entérocolitypes ont été ainsi décrits : entérocolotype sauvage herbivore, entérocolotype domestique herbivore et entérocolotype omnivore, chacun se caractérisant par la prédominance d'un des phylogroupes B2, B1 ou A. Ainsi, les B2 dominants sont retrouvés chez les animaux herbivores sauvages associés à une faible diversité et une densité basse de *E. coli* dans les selles, les B1 dominants chez les herbivores domestiques avec une grande diversité de portage et une quantité intermédiaire de *E. coli* dans les selles et enfin un entérocolotype A dominant chez les omnivores avec une haute densité de *E. coli* dans les selles. L'homme présentait deux types de structures de *E. coli* commensaux : A dominant avec une plus grande diversité de portage de phylogroupes, entérocolotype comparable à celui des animaux omnivores, et B2 dominant et quasi-exclusif. Le phylogroupe B2 n'est exclusif que chez l'homme. Ceci nous a amené à comparer les souches de phylogroupe B2 commensales animales et humaines en fonction de leur capacité à être exclusive dans la population ou pas. Deux gènes *sfa/foc*, codant pour la sous-unité F1C des fimbriae et *clbQ*, de l'îlot génomique *pks* codant pour la génotoxine colibactine et associé au cancer colo-rectal ont montré une association avec le caractère exclusif chez l'homme reflétant une forte capacité de colonisation *in-vivo* spécifique de certaines souches humaines de phylogroupe B2 et créant de ce fait un quatrième entérocolotype spécifiquement humain. L'ensemble de ces données illustre une spécialisation des groupes phylogénétiques en fonction de l'espèce, du régime et du mode de vie avec une particularité chez l'homme expliquée probablement par une forte adaptation de *E. coli* à son hôte.

Ultérieurement, la méthode pourrait être appliquée à l'étude du portage de *E. coli* dans différentes situations physiologiques en fonction de la situation géographique, de l'origine des sujets ou physiopathologiques : hospitalisation, réanimation, effets de l'antibiothérapie, survenue de maladies intercurrentes, infections urinaires à répétitions.

En situation extraintestinale dans des produits pathologiques, l'approche quantitative pourrait être contributive comme facteur prédictif d'infection invasive par exemple : portage de *E. coli* et survenue d'ascite, ascite et choc septique.

Cette méthode pourrait également avoir des applications chez l'animal dans l'élevage, en sciences vétérinaires, et en bactériologie environnementale pour connaître l'origine d'une contamination fécale des eaux.

La mise en évidence lors de ces études de souches commensales humaines chargées en FVs et « super colonisatrices » pose la question suivante : le portage de souches B2 super-colonisatrices est-il prédictif de survenue d'infection invasive à *E. coli* ?

Certes l'impact de la virulence intrinsèque de *E. coli*, pathogène opportuniste, dans la survenue et la gravité de l'infection est moindre que celui de l'état physiologique du sujet, de ses comorbidités et de l'état de son immunité, cependant il demeure un des éléments constitutifs de l'infection.

La deuxième partie de la thèse a abordé le rôle du fond génétique reflété par l'appartenance à différents sous-groupes de B2 dans la virulence médiée par le HPI, ilot génomique codant pour un sidérophore, la yersiniabactine, facteur de virulence et de commensalisme.

Il apparaît que l'impact de la production de la yersiniabactine dépend du fond génétique des souches dans lesquelles il s'exprime. Les souches les plus proches génétiquement telles que C5 et NU14 sont les plus affectées par les mutations du HPI et de ses gènes, et ce quelle que soit leur virulence intrinsèque. L'arrivée d'éléments génétiques par transferts horizontaux peut modifier cet état comme pourrait le faire le plasmide pS88 de la souche S88.

Aussi, le HPI chez *E. coli* qui a été acquis récemment par transfert horizontal en provenance des *Yersiniae*, interagît avec la clonalité de l'espèce qui s'exprime par l'existence des sous-

groupes de B2. Dans les modèles de virulence que nous avons développés, les mêmes gènes présentent, en fonction du fond génétique des différents isolats naturels, des effets différents.

BIBLIOGRAPHIE

- Adiba, S., Nizak, C., van Baalen, M., Denamur, E., and Depaulis, F. (2010) From grazing resistance to pathogenesis: the coincidental evolution of virulence factors. *PloS One* **5**: e11882.
- Alibaud, L., Köhler, T., Coudray, A., Prigent-Combaret, C., Bergeret, E., Perrin, J., et al. (2008) *Pseudomonas aeruginosa* virulence genes identified in a Dictyostelium host model. *Cell. Microbiol.* **10**: 729–740.
- De Almeida, A.M., Guiyoule, A., Guilvout, I., Iteman, I., Baranton, G., and Carniel, E. (1993) Chromosomal *irp2* gene in *Yersinia*: distribution, expression, deletion and impact on virulence. *Microb. Pathog.* **14**: 9–21.
- Alteri, C.J., Hagan, E.C., Sivick, K.E., Smith, S.N., and Mobley, H.L.T. (2009) Mucosal immunization with iron receptor antigens protects against urinary tract infection. *PLoS Pathog.* **5**: e1000586.
- Anderson, M.A., Whitlock, J.E., and Harwood, V.J. (2006) Diversity and distribution of *Escherichia coli* genotypes and antibiotic resistance phenotypes in feces of humans, cattle, and horses. *Appl. Environ. Microbiol.* **72**: 6914–6922.
- Anisimov, R., Brem, D., Heesemann, J., and Rakin, A. (2005) Molecular mechanism of YbtA-mediated transcriptional regulation of divergent overlapping promoters *ybtA* and *irp6* of *Yersinia enterocolitica*. *FEMS Microbiol. Lett.* **250**: 27–32.
- Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., et al. (2011) Enterotypes of the human gut microbiome. *Nature* **473**: 174–180.
- Bachman, M.A., Oyler, J.E., Burns, S.H., Caza, M., Lépine, F., Dozois, C.M., and Weiser, J.N. (2011) *Klebsiella pneumoniae* yersiniabactin promotes respiratory tract infection through evasion of lipocalin 2. *Infect. Immun.* **79**: 3309–3316.
- Bach, S., de Almeida, A., and Carniel, E. (2000) The *Yersinia* high-pathogenicity island is present in different members of the family Enterobacteriaceae. *FEMS Microbiol. Lett.* **183**: 289–294.
- Bach, S., Buchrieser, C., Prentice, M., Guiyoule, A., Msadek, T., and Carniel, E. (1999) The high-pathogenicity island of *Yersinia enterocolitica* Ye8081 undergoes low-frequency deletion but not precise excision, suggesting recent stabilization in the genome. *Infect. Immun.* **67**: 5091–5099.

- Bahrani-Mougeot, F.K., Buckles, E.L., Lockatell, C.V., Hebel, J.R., Johnson, D.E., Tang, C.M., and Donnenberg, M.S. (2002) Type 1 fimbriae and extracellular polysaccharides are preeminent uropathogenic *Escherichia coli* virulence determinants in the murine urinary tract. *Mol. Microbiol.* **45**: 1079–1093.
- Baldy-Chudzik, K., Mackiewicz, P., and Stosik, M. (2008) Phylogenetic background, virulence gene profiles, and genomic diversity in commensal *Escherichia coli* isolated from ten mammal species living in one zoo. *Vet Microbiol* **131**: 173–184.
- Ballyk, M.M., Jones, D.A., and Smith, H.L. (2001) Microbial Competition in Reactors with Wall Attachment. *Microb. Ecol.* **41**: 210–221.
- Bearden, S.W., Fetherston, J.D., and Perry, R.D. (1997) Genetic organization of the yersiniabactin biosynthetic region and construction of avirulent mutants in *Yersinia pestis*. *Infect. Immun.* **65**: 1659–1668.
- Benghezal, M., Fauvarque, M.-O., Tournebize, R., Froquet, R., Marchetti, A., Bergeret, E., et al. (2006) Specific host genes required for the killing of *Klebsiella* bacteria by phagocytes. *Cell. Microbiol.* **8**: 139–148.
- Berche, P., Gaillard, J. L., Simonet, M. Bactériologie: les bactéries des infections humaines. Paris: Flammarion, 1988.
- Berg, R.D. (1996) The indigenous gastrointestinal microflora. *Trends Microbiol* **4**: 430–435.
- Bert, F., Johnson, J.R., Ouattara, B., Leflon-Guibout, V., Johnston, B., Marcon, E., et al. (2010) Genetic diversity and virulence profiles of *Escherichia coli* isolates causing spontaneous bacterial peritonitis and bacteremia in patients with cirrhosis. *J. Clin. Microbiol.* **48**: 2709–2714.
- Berthe, T., Ratajczak, M., Clermont, O., Denamur, E., and Petit, F. (2013) Evidence for coexistence of distinct *Escherichia coli* populations in various aquatic environments and their survival in estuary water. *Appl. Environ. Microbiol.* **79**: 4684–4693.
- Bettelheim, K.A., Faiers, M., and Shooter, R.A. (1972) Serotypes of *Escherichia coli* in normal stools. *Lancet* **2**: 1223–1224.
- Bettelheim, K.A. and Lennox-King, S.M. (1976) The acquisition of *Escherichia coli* by newborn babies. *Infection* **4**: 174–179.
- Bidet, P., Mahjoub-Messai, F., Blanco, J., Blanco, J., Dehem, M., Aujard, Y., et al. (2007) Combined multilocus sequence typing and O serogrouping distinguishes *Escherichia coli* subtypes associated with infant urosepsis and/or meningitis. *J. Infect. Dis.* **196**: 297–303.
- Bingen, E., Picard, B., Brahimi, N., Mathy, S., Desjardins, P., Elion, J., and Denamur, E. (1998) Phylogenetic analysis of *Escherichia coli* strains causing neonatal meningitis suggests

horizontal gene transfer from a predominant pool of highly virulent B2 group strains. *J. Infect. Dis.* **177**: 642–650.

Blum, G., Falbo, V., Caprioli, A., and Hacker, J. (1995) Gene clusters encoding the cytotoxic necrotizing factor type 1, Prs-fimbriae and alpha-hemolysin form the pathogenicity island II of the uropathogenic *Escherichia coli* strain J96. *FEMS Microbiol. Lett.* **126**: 189–195.

Blum, G., Ott, M., Lischewski, A., Ritter, A., Imrich, H., Tschäpe, H., and Hacker, J. (1994) Excision of large DNA regions termed pathogenicity islands from tRNA-specific loci in the chromosome of an *Escherichia coli* wild-type pathogen. *Infect. Immun.* **62**: 606–614.

Bobrov, A.G., Geoffroy, V.A., and Perry, R.D. (2002) Yersiniabactin production requires the thioesterase domain of HMWP2 and YbtD, a putative phosphopantetheinylate transferase. *Infect. Immun.* **70**: 4204–4214.

Bobrov, A.G., Kirillina, O., Fetherston, J.D., Miller, M.C., Burlison, J.A., and Perry, R.D. (2014) The *Yersinia pestis* siderophore, yersiniabactin, and the ZnuABC system both contribute to zinc acquisition and the development of lethal septicaemic plague in mice. *Mol. Microbiol.* **93**: 759–775.

Bonacorsi, S., Clermont, O., Houdouin, V., Cordevant, C., Brahimi, N., Marecat, A., et al. (2003) Molecular analysis and experimental virulence of French and North American *Escherichia coli* neonatal meningitis isolates: identification of a new virulent clone. *J. Infect. Dis.* **187**: 1895–1906.

Bonacorsi, S.P., Clermont, O., Tinsley, C., Le Gall, I., Beaudoin, J.C., Elion, J., et al. (2000) Identification of regions of the *Escherichia coli* chromosome specific for neonatal meningitis-associated strains. *Infect. Immun.* **68**: 2096–2101.

Bonner, J.T., Clarke, W.W., Neely, C.L., and Slifkin, M.K. (1950) The orientation to light and the extremely sensitive orientation to temperature gradients in the slime mold *Dictyostelium discoideum*. *J. Cell. Physiol.* **36**: 149–158.

Bortolussi, R., Ferrieri, P., and Wannamaker, L.W. (1978) Dynamics of *Escherichia coli* infection and meningitis in infant rats. *Infect. Immun.* **22**: 480–485.

Braun, V. and Braun, M. (2002) Iron transport and signaling in *Escherichia coli*. *FEBS Lett.* **529**: 78–85.

Braun, V., Mahren, S., and Sauter, A. (2005) Gene regulation by transmembrane signaling. *Biometals Int. J. Role Met. Ions Biol. Biochem. Med.* **18**: 507–517.

Brem, D., Pelludat, C., Rakin, A., Jacobi, C.A., and Heesemann, J. (2001) Functional analysis of yersiniabactin transport genes of *Yersinia enterocolitica*. *Microbiol. Read. Engl.* **147**: 1115–1127.

- Brenner, D.J., Davis, B.R., Steigerwalt, A.G., Riddle, C.F., McWhorter, A.C., Allen, S.D., et al. (1982) Atypical biogroups of *Escherichia coli* found in clinical specimens and description of *Escherichia hermannii* sp. nov. *J. Clin. Microbiol.* **15**: 703–713.
- Brenner, D.J., Fanning, G.R., Skerman, F.J., and Falkow, S. (1972) Polynucleotide sequence divergence among strains of *Escherichia coli* and closely related organisms. *J. Bacteriol.* **109**: 953–965.
- Brenner, D.J., McWhorter, A.C., Knutson, J.K., and Steigerwalt, A.G. (1982) *Escherichia vulneris*: a new species of *Enterobacteriaceae* associated with human wounds. *J. Clin. Microbiol.* **15**: 1133–1140.
- Britigan, B.E., Rasmussen, G.T., and Cox, C.D. (1997) Augmentation of oxidant injury to human pulmonary epithelial cells by the *Pseudomonas aeruginosa* siderophore pyochelin. *Infect. Immun.* **65**: 1071–1076.
- Britigan, B.E., Rasmussen, G.T., and Cox, C.D. (1994) *Pseudomonas* siderophore pyochelin enhances neutrophil-mediated endothelial cell injury. *Am. J. Physiol.* **266**: L192–198.
- Brock, J.H., Williams, P.H., Licéaga, J., and Wooldridge, K.G. (1991) Relative availability of transferrin-bound iron and cell-derived iron to aerobactin-producing and enterochelin-producing strains of *Escherichia coli* and to other microorganisms. *Infect. Immun.* **59**: 3185–3190.
- Brzuszkiewicz, E., Brüggemann, H., Liesegang, H., Emmerth, M., Olschläger, T., Nagy, G., et al. (2006) How to become a uropathogen: comparative genomic analysis of extraintestinal pathogenic *Escherichia coli* strains. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 12879–12884.
- Buchrieser, C., Brosch, R., Bach, S., Guiyoule, A., and Carniel, E. (1998) The high-pathogenicity island of *Yersinia pseudotuberculosis* can be inserted into any of the three chromosomal asn tRNA genes. *Mol. Microbiol.* **30**: 965–978.
- Buchrieser, C., Rusniok, C., Frangeul, L., Couve, E., Billault, A., Kunst, F., et al. (1999) The 102-kilobase pgm locus of *Yersinia pestis*: sequence analysis and comparison of selected regions among different *Yersinia pestis* and *Yersinia pseudotuberculosis* strains. *Infect. Immun.* **67**: 4851–4861.
- Burgess, N.R., McDermott, S.N., and Whiting, J. (1973) Aerobic bacteria occurring in the hind-gut of the cockroach, *Blatta orientalis*. *J. Hyg. (Lond.)* **71**: 1–7.
- Byappanahalli, M.N., Whitman, R.L., Shively, D.A., Sadowsky, M.J., and Ishii, S. (2006) Population structure, persistence, and seasonality of autochthonous *Escherichia coli* in temperate, coastal forest soil from a Great Lakes watershed. *Environ. Microbiol.* **8**: 504–513.

- Carlos, C., Pires, M.M., Stoppe, N.C., Hachich, E.M., Sato, M.I.Z., Gomes, T.A.T., et al. (2010) *Escherichia coli* phylogenetic group determination and its application in the identification of the major animal source of fecal contamination. *BMC Microbiol.* **10**: 161.
- Carniel, E. (1999) The *Yersinia* high-pathogenicity island. *Int. Microbiol. Off. J. Span. Soc. Microbiol.* **2**: 161–167.
- Carniel, E. (2001) The *Yersinia* high-pathogenicity island: an iron-uptake island. *Microbes Infect. Inst. Pasteur* **3**: 561–569.
- Carniel, E., Guilvout, I., and Prentice, M. (1996) Characterization of a large chromosomal “high-pathogenicity island” in biotype 1B *Yersinia enterocolitica*. *J. Bacteriol.* **178**: 6743–6751.
- Caugant, D.A., Levin, B.R., and Selander, R.K. (1984) Distribution of multilocus genotypes of *Escherichia coli* within and between host families. *J. Hyg. (Lond.)* **92**: 377–384.
- Caugant, D.A., Levin, B.R., and Selander, R.K. (1981) Genetic diversity and temporal variation in the *E. coli* population of a human host. *Genetics* **98**: 467–490.
- Chaturvedi, K.S., Hung, C.S., Crowley, J.R., Stapleton, A.E., and Henderson, J.P. (2012) The siderophore yersiniabactin binds copper to protect pathogens during infection. *Nat. Chem. Biol.* **8**: 731–736.
- Chaturvedi, K.S., Hung, C.S., Giblin, D.E., Urushidani, S., Austin, A.M., Dinauer, M.C., and Henderson, J.P. (2014) Cupric yersiniabactin is a virulence-associated superoxide dismutase mimic. *ACS Chem. Biol.* **9**: 551–561.
- Chaudhuri, R.R. and Henderson, I.R. (2012) The evolution of the *Escherichia coli* phylogeny. *Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis.* **12**: 214–226.
- Chen, G., Zhuchenko, O., and Kuspa, A. (2007) Immune-like phagocyte activity in the social amoeba. *Science* **317**: 678–681.
- Cirl, C., Wieser, A., Yadav, M., Duerr, S., Schubert, S., Fischer, H., et al. (2008) Subversion of Toll-like receptor signaling by a unique family of bacterial Toll/interleukin-1 receptor domain-containing proteins. *Nat. Med.* **14**: 399–406.
- Clermont, O. (2008) Evidence for a human-specific *Escherichia coli* clone. *Env. Microbiol* **10**: 1000–1006.
- Clermont, O., Bonacorsi, S., and Bingen, E. (2001) The *Yersinia* high-pathogenicity island is highly predominant in virulence-associated phylogenetic groups of *Escherichia coli*. *FEMS Microbiol. Lett.* **196**: 153–157.

- Clermont, O., Christenson, J.K., Daubié, A.-S., Gordon, D.M., and Denamur, E. (2014) Development of an allele-specific PCR for *Escherichia coli* B2 sub-typing, a rapid and easy to perform substitute of multilocus sequence typing. *J. Microbiol. Methods* **101**: 24–27.
- Clermont, O., Christenson, J.K., Denamur, E., and Gordon, D.M. (2013) The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ. Microbiol. Rep.* **5**: 58–65.
- Clermont, O., Lavollay, M., Vimont, S., Deschamps, C., Forestier, C., Branger, C., et al. (2008) The CTX-M-15-producing *Escherichia coli* diffusing clone belongs to a highly virulent B2 phylogenetic subgroup. *J. Antimicrob. Chemother.* **61**: 1024–1028.
- Clermont, O., Olier, M., Hoede, C., Diancourt, L., Brisse, S., Keroudean, M., et al. (2011) Animal and human pathogenic *Escherichia coli* strains share common genetic backgrounds. *Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis.* **11**: 654–662.
- Coffman, T.J., Cox, C.D., Edeker, B.L., and Britigan, B.E. (1990) Possible role of bacterial siderophores in inflammation. Iron bound to the *Pseudomonas* siderophore pyochelin can function as a hydroxyl radical catalyst. *J. Clin. Invest.* **86**: 1030–1037.
- Conway, T., Krogfelt, K.A., and Cohen, P.S. (2004) EcoSal.
- Cooke, E.M., Hettiaratchy, I.G., and Buck, A.C. (1972) Fate of ingested *Escherichia coli* in normal persons. *J Med Microbiol* **5**: 361–369.
- Cosson, P. and Soldati, T. (2008) Eat, kill or die: when amoeba meets bacteria. *Curr. Opin. Microbiol.* **11**: 271–276.
- Croxen, M.A. and Finlay, B.B. (2010) Molecular mechanisms of *Escherichia coli* pathogenicity. *Nat. Rev. Microbiol.* **8**: 26–38.
- Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97**: 6640–6645.
- Demir, M. and Kaleli, I. (2004) Production by *Escherichia coli* isolates of siderophore and other virulence factors and their pathogenic role in a cutaneous infection model. *Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis.* **10**: 1011–1014.
- Deschamps, C. (2009) Multiple acquisitions of CTX-M plasmids in the rare D2 genotype of *Escherichia coli* provide evidence for convergent evolution. *Microbiology* **155**: 1656–1668.
- Desjardins, P., Picard, B., Kaltenbock, B., Elion, J., and Denamur, E. (1995) Sex in *Escherichia coli* does not disrupt the clonal structure of the population: evidence from random amplified polymorphic DNA and restriction-fragment-length polymorphism. *J Mol Evol* **41**: 440–448.

- Diard, M. (2007) *Caenorhabditis elegans* as a simple model to study phenotypic and genetic virulence determinants of extraintestinal pathogenic *Escherichia coli*. *Microbes Infect* **9**: 214–223.
- Diard, M., Garry, L., Selva, M., Mosser, T., Denamur, E., and Matic, I. (2010) Pathogenicity-associated islands in extraintestinal pathogenic *Escherichia coli* are fitness elements involved in intestinal colonization. *J. Bacteriol.* **192**: 4885–4893.
- Dixit, S.M. (2004) Diversity analysis of commensal porcine *Escherichia coli* [mdash] associations between genotypes and habitat in the porcine gastrointestinal tract. *Microbiology* **150**: 1735–1740.
- Dobrindt, U., Blum-Oehler, G., Nagy, G., Schneider, G., Johann, A., Gottschalk, G., and Hacker, J. (2002) Genetic structure and distribution of four pathogenicity islands (PAI I(536) to PAI IV(536)) of uropathogenic *Escherichia coli* strain 536. *Infect. Immun.* **70**: 6365–6372.
- Dobrindt, U., Hochhut, B., Hentschel, U., and Hacker, J. (2004) Genomic islands in pathogenic and environmental microorganisms. *Nat. Rev. Microbiol.* **2**: 414–424.
- Durant, L., Metais, A., Soulama-Mouze, C., Genevard, J.-M., Nassif, X., and Escaich, S. (2007) Identification of candidates for a subunit vaccine against extraintestinal pathogenic *Escherichia coli*. *Infect. Immun.* **75**: 1916–1925.
- Duriez, P., Clermont, O., Bonacorsi, S., Bingen, E., Chaventré, A., Elion, J., et al. (2001) Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. *Microbiol. Read. Engl.* **147**: 1671–1676.
- Escobar-Paramo, P. (2004) Decreasing the effects of horizontal gene transfer on bacterial phylogeny: the *Escherichia coli* case study. *Mol Phylogenet Evol* **30**: 243–250.
- Escobar-Páramo, P., Clermont, O., Blanc-Potard, A.-B., Bui, H., Le Bouguénec, C., and Denamur, E. (2004) A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. *Mol. Biol. Evol.* **21**: 1085–1094.
- Escobar-Páramo, P., Grenet, K., Le Menac’h, A., Rode, L., Salgado, E., Amorin, C., et al. (2004) Large-scale population structure of human commensal *Escherichia coli* isolates. *Appl. Environ. Microbiol.* **70**: 5698–5700.
- Escobar-Páramo, P., Le Menac’h, A., Le Gall, T., Amorin, C., Gouriou, S., Picard, B., et al. (2006) Identification of forces shaping the commensal *Escherichia coli* genetic structure by comparing animal and human isolates. *Environ. Microbiol.* **8**: 1975–1984.
- Fang, C.-T., Chuang, Y.-P., Shun, C.-T., Chang, S.-C., and Wang, J.-T. (2004) A novel virulence gene in *Klebsiella pneumoniae* strains causing primary liver abscess and septic metastatic complications. *J. Exp. Med.* **199**: 697–705.

- Fantin, B., Duval, X., Massias, L., Alavoine, L., Chau, F., Retout, S., et al. (2009) Ciprofloxacin dosage and emergence of resistance in human commensal bacteria. *J. Infect. Dis.* **200**: 390–398.
- Farmer, J.J., 3rd, Fanning, G.R., Davis, B.R., O'Hara, C.M., Riddle, C., Hickman-Brenner, F.W., et al. (1985) *Escherichia fergusonii* and *Enterobacter taylorae*, two new species of Enterobacteriaceae isolated from clinical specimens. *J. Clin. Microbiol.* **21**: 77–81.
- Feldmann, F., Sorsa, L.J., Hildinger, K., and Schubert, S. (2007) The salmochelin siderophore receptor IroN contributes to invasion of urothelial cells by extraintestinal pathogenic *Escherichia coli* in vitro. *Infect. Immun.* **75**: 3183–3187.
- Fetherston, J.D., Bearden, S.W., and Perry, R.D. (1996) YbtA, an AraC-type regulator of the *Yersinia pestis* pesticin/yersiniabactin receptor. *Mol. Microbiol.* **22**: 315–325.
- Fetherston, J.D., Bertolino, V.J., and Perry, R.D. (1999) YbtP and YbtQ: two ABC transporters required for iron uptake in *Yersinia pestis*. *Mol. Microbiol.* **32**: 289–299.
- Fetherston, J.D., Kirillina, O., Bobrov, A.G., Paulley, J.T., and Perry, R.D. (2010) The yersiniabactin transport system is critical for the pathogenesis of bubonic and pneumonic plague. *Infect. Immun.* **78**: 2045–2052.
- Fetherston, J.D., Lillard, J.W., and Perry, R.D. (1995) Analysis of the pesticin receptor from *Yersinia pestis*: role in iron-deficient growth and possible regulation by its siderophore. *J. Bacteriol.* **177**: 1824–1833.
- Fetherston, J.D. and Perry, R.D. (1994) The pigmentation locus of *Yersinia pestis* KIM6+ is flanked by an insertion sequence and includes the structural genes for pesticin sensitivity and HMWP2. *Mol. Microbiol.* **13**: 697–708.
- Fischbach, M.A., Lin, H., Zhou, L., Yu, Y., Abergel, R.J., Liu, D.R., et al. (2006) The pathogen-associated iroA gene cluster mediates bacterial evasion of lipocalin 2. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 16502–16507.
- Flo, T.H., Smith, K.D., Sato, S., Rodriguez, D.J., Holmes, M.A., Strong, R.K., et al. (2004) Lipocalin 2 mediates an innate immune response to bacterial infection by sequestering iron. *Nature* **432**: 917–921.
- Foxman, B. and Frerichs, R.R. (1985) Epidemiology of urinary tract infection: II. Diet, clothing, and urination habits. *Am. J. Public Health* **75**: 1314–1317.
- Freter, R., Brickner, H., Fekete, J., Vickerman, M.M., and Carey, K.E. (1983) Survival and implantation of *Escherichia coli* in the intestinal tract. *Infect Immun* **39**: 686–703.

- Froquet, R., Cherix, N., Burr, S.E., Frey, J., Vilches, S., Tomas, J.M., and Cosson, P. (2007) Alternative host model to evaluate *Aeromonas* virulence. *Appl. Environ. Microbiol.* **73**: 5657–5659.
- Furet, J.-P., Firmesse, O., Gourmelon, M., Bridonneau, C., Tap, J., Mondot, S., et al. (2009) Comparative assessment of human and farm animal faecal microbiota using real-time quantitative PCR. *FEMS Microbiol. Ecol.* **68**: 351–362.
- Le Gall, T., Clermont, O., Gouriou, S., Picard, B., Nassif, X., Denamur, E., and Tenaillon, O. (2007) Extraintestinal virulence is a coincidental by-product of commensalism in B2 phylogenetic group *Escherichia coli* strains. *Mol. Biol. Evol.* **24**: 2373–2384.
- Garcia, E.C., Brumbaugh, A.R., and Mobley, H.L.T. (2011) Redundancy and specificity of *Escherichia coli* iron acquisition systems during urinary tract infection. *Infect. Immun.* **79**: 1225–1235.
- Gehring, A.M., DeMoll, E., Fetherston, J.D., Mori, I., Mayhew, G.F., Blattner, F.R., et al. (1998) Iron acquisition in plague: modular logic in enzymatic biogenesis of yersiniabactin by *Yersinia pestis*. *Chem. Biol.* **5**: 573–586.
- Geoffroy, V.A., Fetherston, J.D., and Perry, R.D. (2000) *Yersinia pestis* YbtU and YbtT are involved in synthesis of the siderophore yersiniabactin but have different effects on regulation. *Infect. Immun.* **68**: 4452–4461.
- Gibreel, T.M., Dodgson, A.R., Cheesbrough, J., Fox, A.J., Bolton, F.J., and Upton, M. (2012) Population structure, virulence potential and antibiotic susceptibility of uropathogenic *Escherichia coli* from Northwest England. *J. Antimicrob. Chemother.* **67**: 346–356.
- Gordon, D.M., Bauer, S., and Johnson, J.R. (2002) The genetic structure of *Escherichia coli* populations in primary and secondary habitats. *Microbiol. Read. Engl.* **148**: 1513–1522.
- Gordon, D.M. and Cowling, A. (2003) The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology* **149**: 3575–3586.
- Gordon, D.M., Stern, S.E., and Collignon, P.J. (2005) Influence of the age and sex of human hosts on the distribution of *Escherichia coli* ECOR groups and virulence traits. *Microbiol. Read. Engl.* **151**: 15–23.
- Goulet, P. and Picard, B. (1989) Comparative electrophoretic polymorphism of esterases and other enzymes in *Escherichia coli*. *J Gen Microbiol* **135**: 135–143.
- Goulet, P. and Picard, B. (1986) Comparative esterase electrophoretic polymorphism of *Escherichia coli* isolates obtained from animal and human sources. *J Gen Microbiol* **132**: 1843–1851.

- Grimont, P.A. (1988) Use of DNA reassociation in bacterial classification. *Can. J. Microbiol.* **34**: 541–546.
- Haag, H., Hantke, K., Drechsel, H., Stojiljkovic, I., Jung, G., and Zähler, H. (1993) Purification of yersiniabactin: a siderophore and possible virulence factor of *Yersinia enterocolitica*. *J. Gen. Microbiol.* **139**: 2159–2165.
- Hacker, J., Bender, L., Ott, M., Wingender, J., Lund, B., Marre, R., and Goebel, W. (1990) Deletions of chromosomal regions coding for fimbriae and hemolysins occur in vitro and in vivo in various extraintestinal *Escherichia coli* isolates. *Microb. Pathog.* **8**: 213–225.
- Hacker, J., Blum-Oehler, G., Mühldorfer, I., and Tschäpe, H. (1997) Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol. Microbiol.* **23**: 1089–1097.
- Hacker, J., Hughes, C., Hof, H., and Goebel, W. (1983) Cloned hemolysin genes from *Escherichia coli* that cause urinary tract infection determine different levels of toxicity in mice. *Infect. Immun.* **42**: 57–63.
- Hacker, J. and Kaper, J.B. (2000) Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* **54**: 641–679.
- Hacker, J., Knapp, S., and Goebel, W. (1983) Spontaneous deletions and flanking regions of the chromosomally inherited hemolysin determinant of an *Escherichia coli* O6 strain. *J. Bacteriol.* **154**: 1145–1152.
- Hagan, E.C., Lloyd, A.L., Rasko, D.A., Faerber, G.J., and Mobley, H.L.T. (2010) *Escherichia coli* global gene expression in urine from women with urinary tract infection. *PLoS Pathog.* **6**: e1001187.
- Hagan, E.C. and Mobley, H.L.T. (2007) Uropathogenic *Escherichia coli* outer membrane antigens expressed during urinary tract infection. *Infect. Immun.* **75**: 3941–3949.
- Hancock, V., Ferrières, L., and Klemm, P. (2008) The ferric yersiniabactin uptake receptor FyuA is required for efficient biofilm formation by urinary tract infectious *Escherichia coli* in human urine. *Microbiol. Read. Engl.* **154**: 167–175.
- Hancock, V. and Klemm, P. (2007) Global gene expression profiling of asymptomatic bacteriuria *Escherichia coli* during biofilm growth in human urine. *Infect. Immun.* **75**: 966–976.
- Hartl, D.L. (1992) Population genetics of microbial organisms. *Curr. Opin. Genet. Dev.* **2**: 937–942.
- Heesemann, J., Hantke, K., Vocke, T., Saken, E., Rakin, A., Stojiljkovic, I., and Berner, R. (1993) Virulence of *Yersinia enterocolitica* is closely associated with siderophore production,

expression of an iron-repressible outer membrane polypeptide of 65,000 Da and pesticin sensitivity. *Mol. Microbiol.* **8**: 397–408.

Henderson, J.P., Crowley, J.R., Pinkner, J.S., Walker, J.N., Tsukayama, P., Stamm, W.E., et al. (2009) Quantitative metabolomics reveals an epigenetic blueprint for iron acquisition in uropathogenic *Escherichia coli*. *PLoS Pathog.* **5**: e1000305.

Hershberg, R., Tang, H., and Petrov, D.A. (2007) Reduced selection leads to accelerated gene loss in *Shigella*. *Genome Biol.* **8**: R164.

Herzer, P.J., Inouye, S., Inouye, M., and Whittam, T.S. (1990) Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. *J. Bacteriol.* **172**: 6175–6181.

Hilbi, H., Weber, S.S., Ragaz, C., Nyfeler, Y., and Urwyler, S. (2007) Environmental predators as models for bacterial pathogenesis. *Environ. Microbiol.* **9**: 563–575.

Hileti, D., Panayiotidis, P., and Hoffbrand, A.V. (1995) Iron chelators induce apoptosis in proliferating cells. *Br. J. Haematol.* **89**: 181–187.

Hobman, J.L., Penn, C.W., and Pallen, M.J. (2007) Laboratory strains of *Escherichia coli*: model citizens or deceitful delinquents growing old disgracefully? *Mol Microbiol* **64**: 881–885.

Hudault, S., Guignot, J., and Servin, A.L. (2001) *Escherichia coli* strains colonising the gastrointestinal tract protect germfree mice against *Salmonella typhimurium* infection. *Gut* **49**: 47–55.

Huys, G., Cnockaert, M., Janda, J.M., and Swings, J. (2003) *Escherichia albertii* sp. nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. *Int. J. Syst. Evol. Microbiol.* **53**: 807–810.

Hyma, K.E., Lacher, D.W., Nelson, A.M., Bumbaugh, A.C., Janda, J.M., Strockbine, N.A., et al. (2005) Evolutionary genetics of a new pathogenic *Escherichia* species: *Escherichia albertii* and related *Shigella boydii* strains. *J. Bacteriol.* **187**: 619–628.

Jacobi, C.A., Gregor, S., Rakin, A., and Heesemann, J. (2001) Expression analysis of the yersiniabactin receptor gene *fyuA* and the heme receptor *hemR* of *Yersinia enterocolitica* in vitro and *in vivo* using the reporter genes for green fluorescent protein and luciferase. *Infect. Immun.* **69**: 7772–7782.

Jauregui, F. (2004) Effects of intrapartum penicillin prophylaxis on intestinal bacterial colonization in infants. *J Clin Microbiol* **42**: 5184–5188.

- Jauregui, F., Landraud, L., Passet, V., Diancourt, L., Frapy, E., Guigon, G., et al. (2008) Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains. *BMC Genomics* **9**: 560.
- Johnson, J.R. (1991) Virulence factors in *Escherichia coli* urinary tract infection. *Clin. Microbiol. Rev.* **4**: 80–128.
- Johnson, J.R., Clabots, C., and Kuskowski, M.A. (2008) Multiple-host sharing, long-term persistence, and virulence of *Escherichia coli* clones from human and animal household members. *J Clin Microbiol* **46**: 4078–4082.
- Johnson, J.R., Delavari, P., Kuskowski, M., and Stell, A.L. (2001) Phylogenetic distribution of extraintestinal virulence-associated traits in *Escherichia coli*. *J. Infect. Dis.* **183**: 78–88.
- Johnson, J.R., Jelacic, S., Schoening, L.M., Clabots, C., Shaikh, N., Mobley, H.L.T., and Tarr, P.I. (2005) The IrgA homologue adhesin Iha is an *Escherichia coli* virulence factor in murine urinary tract infection. *Infect. Immun.* **73**: 965–971.
- Johnson, J.R., Johnston, B., Clabots, C.R., Kuskowski, M.A., Roberts, E., and DebRoy, C. (2007) Virulence Genotypes and Phylogenetic Background of *Escherichia coli* Serogroup O6 Isolates from Humans, Dogs, and Cats. *J. Clin. Microbiol.* **46**: 417–422.
- Johnson, J.R., Owens, K., Gajewski, A., and Clabots, C. (2008) *Escherichia coli* colonization patterns among human household members and pets, with attention to acute urinary tract infection. *J Infect Dis* **197**: 218–224.
- Johnson, J.R., Owens, K.L., Clabots, C.R., Weissman, S.J., and Cannon, S.B. (2006) Phylogenetic relationships among clonal groups of extraintestinal pathogenic *Escherichia coli* as assessed by multi-locus sequence analysis. *Microbes Infect* **8**: 1702–1713.
- Johnson, J.R. and Russo, T.A. (2005) Molecular epidemiology of extraintestinal pathogenic (uropathogenic) *Escherichia coli*. *Int. J. Med. Microbiol. IJMM* **295**: 383–404.
- Johnson, J.R. and Stell, A.L. (2000) Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *J. Infect. Dis.* **181**: 261–272.
- Johnson, J.R., Weissman, S.J., Stell, A.L., Trintchina, E., Dykhuizen, D.E., and Sokurenko, E.V. (2001) Clonal and Pathotypic Analysis of Archetypal *Escherichia coli* Cystitis Isolate NU14. *J. Infect. Dis.* **184**: 1556–1565.
- Kanamaru, S., Kurazono, H., Ishitoya, S., Terai, A., Habuchi, T., Nakano, M., et al. (2003) Distribution and genetic association of putative uropathogenic virulence factors *iroN*, *iha*, *kpsMT*, *ompT* and *usp* in *Escherichia coli* isolated from urinary tract infections in Japan. *J. Urol.* **170**: 2490–2493.

- Kaper, J.B., Nataro, J.P., and Mobley, H.L. (2004) Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.* **2**: 123–140.
- Karch, H., Schubert, S., Zhang, D., Zhang, W., Schmidt, H., Olschläger, T., and Hacker, J. (1999) A genomic island, termed high-pathogenicity island, is present in certain non-O157 Shiga toxin-producing *Escherichia coli* clonal lineages. *Infect. Immun.* **67**: 5994–6001.
- Kirillina, O., Bobrov, A.G., Fetherston, J.D., and Perry, R.D. (2006) Hierarchy of iron uptake systems: Yfu and Yiu are functional in *Yersinia pestis*. *Infect. Immun.* **74**: 6171–6178.
- Konopka, K. and Neilands, J.B. (1984) Effect of serum albumin on siderophore-mediated utilization of transferrin iron. *Biochemistry (Mosc.)* **23**: 2122–2127.
- Kosek, M., Bern, C., and Guerrant, R.L. (2003) The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bull World Health Organ* **81**: 197–204.
- Lamont, I.L., Beare, P.A., Ochsner, U., Vasil, A.I., and Vasil, M.L. (2002) Siderophore-mediated signaling regulates virulence factor production in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U. S. A.* **99**: 7072–7077.
- Lautenbach, E., Bilker, W.B., Tolomeo, P., and Maslow, J.N. (2008) Impact of diversity of colonizing strains on strategies for sampling *Escherichia coli* from fecal specimens. *J. Clin. Microbiol.* **46**: 3094–3096.
- Lawlor, M.S., O’connor, C., and Miller, V.L. (2007) Yersiniabactin is a virulence factor for *Klebsiella pneumoniae* during pulmonary infection. *Infect. Immun.* **75**: 1463–1472.
- Lawrence, J.G., Ochman, H., and Hartl, D.L. (1991) Molecular and evolutionary relationships among enteric bacteria. *J Gen Microbiol* **137**: 1911–1921.
- Lecointre, G., Rachdi, L., Darlu, P., and Denamur, E. (1998) *Escherichia coli* molecular phylogeny using the incongruence length difference test. *Mol Biol Evol* **15**: 1685–1695.
- Lefort, A., Panhard, X., Clermont, O., Woerther, P.-L., Branger, C., Mentré, F., et al. (2011) Host Factors and Portal of Entry Outweigh Bacterial Determinants To Predict the Severity of *Escherichia coli* Bacteremia. *J. Clin. Microbiol.* **49**: 777–783.
- Lescat, M. (2009) aes, the gene encoding the esterase B in *Escherichia coli*, is a powerful phylogenetic marker of the species. *BMC Microbiol* **9**: 273.
- Lescat, M., Clermont, O., Woerther, P.L., Glodt, J., Dion, S., Skurnik, D., et al. (2013) Commensal *Escherichia coli* strains in Guiana reveal a high genetic diversity with host-dependant population structure. *Environ. Microbiol. Rep.* **5**: 49–57.
- Lesic, B. and Carniel, E. (2005) Horizontal transfer of the high-pathogenicity island of *Yersinia pseudotuberculosis*. *J. Bacteriol.* **187**: 3352–3358.

- Levert, M., Zamfir, O., Clermont, O., Bouvet, O., Lespinats, S., Hipeaux, M.C., et al. (2010) Molecular and evolutionary bases of within-patient genotypic and phenotypic diversity in *Escherichia coli* extraintestinal infections. *PLoS Pathog.* **6**: e1001125.
- Lidin-Janson, G., Kaijser, B., Lincoln, K., Olling, S., and Wedel, H. (1978) The homogeneity of the faecal coliform flora of normal school-girls, characterized by serological and biochemical properties. *Med. Microbiol. Immunol. (Berl.)* **164**: 247–253.
- Lima, W.C., Lelong, E., and Cosson, P. (2011) What can *Dictyostelium* bring to the study of *Pseudomonas* infections? *Semin. Cell Dev. Biol.* **22**: 77–81.
- Livermore, D.M. and Woodford, N. (2006) The beta-lactamase threat in Enterobacteriaceae, *Pseudomonas* and *Acinetobacter*. *Trends Microbiol.* **14**: 413–420.
- Lloyd, A.L., Henderson, T.A., Vigil, P.D., and Mobley, H.L.T. (2009) Genomic islands of uropathogenic *Escherichia coli* contribute to virulence. *J. Bacteriol.* **191**: 3469–3481.
- De Louvois, J. (1994) Acute bacterial meningitis in the newborn. *J. Antimicrob. Chemother.* **34 Suppl A**: 61–73.
- Low, D., David, V., Lark, D., Schoolnik, G., and Falkow, S. (1984) Gene clusters governing the production of hemolysin and mannose-resistant hemagglutination are closely linked in *Escherichia coli* serotype O4 and O6 isolates from urinary tract infections. *Infect. Immun.* **43**: 353–358.
- Luo, C., Walk, S.T., Gordon, D.M., Feldgarden, M., Tiedje, J.M., and Konstantinidis, K.T. (2011) Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 7200–7205.
- Lv, H. and Henderson, J.P. (2011) Yersinia high pathogenicity island genes modify the *Escherichia coli* primary metabolome independently of siderophore production. *J. Proteome Res.* **10**: 5547–5554.
- Mahjoub-Messai, F., Bidet, P., Caro, V., Diancourt, L., Biran, V., Aujard, Y., et al. (2011) *Escherichia coli* isolates causing bacteremia via gut translocation and urinary tract infection in young infants exhibit different virulence genotypes. *J. Infect. Dis.* **203**: 1844–1849.
- March, C., Cano, V., Moranta, D., Llobet, E., Pérez-Gutiérrez, C., Tomás, J.M., et al. (2013) Role of bacterial surface structures on the interaction of *Klebsiella pneumoniae* with phagocytes. *PloS One* **8**: e56847.
- Marre, R., Hacker, J., Henkel, W., and Goebel, W. (1986) Contribution of cloned virulence factors from uropathogenic *Escherichia coli* strains to nephropathogenicity in an experimental rat pyelonephritis model. *Infect. Immun.* **54**: 761–767.

- Maurelli, A.T., Fernández, R.E., Bloch, C.A., Rode, C.K., and Fasano, A. (1998) “Black holes” and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **95**: 3943–3948.
- Mayr, E. (1963). *Animal species and evolution*. Harvard University Press, Cambridge, Massachussets.
- McDaniel, T.K., Jarvis, K.G., Donnenberg, M.S., and Kaper, J.B. (1995) A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc. Natl. Acad. Sci. U. S. A.* **92**: 1664–1668.
- Middendorf, B., Hochhut, B., Leipold, K., Dobrindt, U., Blum-Oehler, G., and Hacker, J. (2004) Instability of pathogenicity islands in uropathogenic *Escherichia coli* 536. *J. Bacteriol.* **186**: 3086–3096.
- Milkman, R. (1973) Electrophoretic variation in *Escherichia coli* from natural sources. *Science* **182**: 1024–1026.
- Miller, M.C., Parkin, S., Fetherston, J.D., Perry, R.D., and Demoll, E. (2006) Crystal structure of ferric-yersiniabactin, a virulence factor of *Yersinia pestis*. *J. Inorg. Biochem.* **100**: 1495–1500.
- Mitsumori, K., Terai, A., Yamamoto, S., and Yoshida, O. (1998) Identification of S, F1C and three PapG fimbrial adhesins in uropathogenic *Escherichia coli* by polymerase chain reaction. *FEMS Immunol. Med. Microbiol.* **21**: 261–268.
- Mitsuoka, T. and Hayakawa, K. (1972) The fecal flora of man. I. Communication: the composition of the fecal flora of different age groups. *Zentralbl Bakteriol Orig A* **223**: 333–342.
- Mobley, H.L., Green, D.M., Trifillis, A.L., Johnson, D.E., Chippendale, G.R., Lockatell, C.V., et al. (1990) Pyelonephritogenic *Escherichia coli* and killing of cultured human renal proximal tubular epithelial cells: role of hemolysin in some strains. *Infect. Immun.* **58**: 1281–1289.
- Moissenet, D., Salauze, B., Clermont, O., Bingen, E., Arlet, G., Denamur, E., et al. (2010) Meningitis caused by *Escherichia coli* producing TEM-52 extended-spectrum beta-lactamase within an extensive outbreak in a neonatal ward: epidemiological investigation and characterization of the strain. *J. Clin. Microbiol.* **48**: 2459–2463.
- Mokracka, J., Koczura, R., and Kaznowski, A. (2004) Yersiniabactin and other siderophores produced by clinical isolates of *Enterobacter* spp. and *Citrobacter* spp. *FEMS Immunol. Med. Microbiol.* **40**: 51–55.

- Moreno, E., Andreu, A., Pigrau, C., Kuskowski, M.A., Johnson, J.R., and Prats, G. (2008) Relationship between *Escherichia coli* strains causing acute cystitis in women and the fecal *E. coli* population of the host. *J. Clin. Microbiol.* **46**: 2529–2534.
- Moreno, E., Johnson, J.R., Pérez, T., Prats, G., Kuskowski, M.A., and Andreu, A. (2009) Structure and urovirulence characteristics of the fecal *Escherichia coli* population among healthy women. *Microbes Infect. Inst. Pasteur* **11**: 274–280.
- Mulvey, M.A., Lopez-Boado, Y.S., Wilson, C.L., Roth, R., Parks, W.C., Heuser, J., and Hultgren, S.J. (1998) Induction and evasion of host defenses by type 1-piliated uropathogenic *Escherichia coli*. *Science* **282**: 1494–1497.
- Myhal, M.L., Laux, D.C., and Cohen, P.S. (1982) Relative colonizing abilities of human fecal and K 12 strains of *Escherichia coli* in the large intestines of streptomycin-treated mice. *Eur J Clin Microbiol* **1**: 186–192.
- Nataro, J.P. and Kaper, J.B. (1998) Diarrheagenic *Escherichia coli*. *Clin. Microbiol. Rev.* **11**: 142–201.
- Nègre, V.L., Bonacorsi, S., Schubert, S., Bidet, P., Nassif, X., and Bingen, E. (2004) The siderophore receptor IroN, but not the high-pathogenicity island or the hemin receptor ChuA, contributes to the bacteremic step of *Escherichia coli* neonatal meningitis. *Infect. Immun.* **72**: 1216–1220.
- Neidhardt, F. C., Curtiss III, R., Ingraham, J. L. *et al.* (1996). *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. ASM Press, Washington.
- Nelson, A.L., Barasch, J.M., Bunte, R.M., and Weiser, J.N. (2005) Bacterial colonization of nasal mucosa induces expression of siderocalin, an iron-sequestering component of innate immunity. *Cell. Microbiol.* **7**: 1404–1417.
- Nicolas-Chanoine, M.-H., Blanco, J., Leflon-Guibout, V., Demarty, R., Alonso, M.P., Caniça, M.M., *et al.* (2008) Intercontinental emergence of *Escherichia coli* clone O25:H4-ST131 producing CTX-M-15. *J. Antimicrob. Chemother.* **61**: 273–281.
- Nowrouzian, F. (2003) *Escherichia coli* in infants' intestinal microflora: colonization rate, strain turnover, and virulence gene carriage. *Pediatr Res* **54**: 8–14.
- Nowrouzian, F.L., Ostblom, A.E., Wold, A.E., and Adlerberth, I. (2009) Phylogenetic group B2 *Escherichia coli* strains from the bowel microbiota of Pakistani infants carry few virulence genes and lack the capacity for long-term persistence. *Clin Microbiol Infect* **15**: 466–472.
- Nowrouzian, F.L., Wold, A.E., and Adlerberth, I. (2005) *Escherichia coli* strains belonging to phylogenetic group B2 have superior capacity to persist in the intestinal microflora of infants. *J. Infect. Dis.* **191**: 1078–1083.

- Obata-Yasuoka, M., Ba-Thein, W., Tsukamoto, T., Yoshikawa, H., and Hayashi, H. (2002) Vaginal *Escherichia coli* share common virulence factor profiles, serotypes and phylogeny with other extraintestinal *E. coli*. *Microbiology* **148**: 2745–2752.
- O’Brien, I.G. and Gibson, F. (1970) The structure of enterochelin and related 2,3-dihydroxy-N-benzoylserine conjugates from *Escherichia coli*. *Biochim. Biophys. Acta* **215**: 393–402.
- Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299–304.
- Ochman, H. and Selander, R.K. (1984) Standard reference strains of *Escherichia coli* from natural populations. *J Bacteriol* **157**: 690–693.
- Orskov, F. and Orskov, I. (1992) *Escherichia coli* serotyping and disease in man and animals. *Can J Microbiol* **38**: 699–704.
- Orskov, I., Orskov, F., Birch-Andersen, A., Kanamori, M., and Svanborg-Edén, C. (1982) O, K, H and fimbrial antigens in *Escherichia coli* serotypes associated with pyelonephritis and cystitis. *Scand. J. Infect. Dis. Suppl.* **33**: 18–25.
- Otto, G.P. and Kessin, R.H. (2001) The intriguing biology of *Dictyostelium discoideum*. Meeting report: International Dictyostelium Conference 2001. *Protist* **152**: 243–248.
- Paauw, A., Leverstein-van Hall, M.A., van Kessel, K.P.M., Verhoef, J., and Fluit, A.C. (2009) Yersiniabactin reduces the respiratory oxidative stress response of innate immune cells. *PLoS One* **4**: e8240.
- Pallecchi, L. (2007) Population structure and resistance genes in antibiotic-resistant bacteria from a remote community with minimal antibiotic exposure. *Antimicrob Agents Chemother* **51**: 1179–1184.
- Paradis, S., Boissinot, M., Paquette, N., Bélanger, S.D., Martel, E.A., Boudreau, D.K., et al. (2005) Phylogeny of the Enterobacteriaceae based on genes encoding elongation factor Tu and F-ATPase beta-subunit. *Int. J. Syst. Evol. Microbiol.* **55**: 2013–2025.
- Peigne, C., Bidet, P., Mahjoub-Messai, F., Plainvert, C., Barbe, V., Médigue, C., et al. (2009) The plasmid of *Escherichia coli* strain S88 (O45:K1:H7) that causes neonatal meningitis is closely related to avian pathogenic *E. coli* plasmids and is associated with high-level bacteremia in a neonatal rat meningitis model. *Infect. Immun.* **77**: 2272–2284.
- Pelludat, C., Brem, D., and Heesemann, J. (2003) Irp9, encoded by the high-pathogenicity island of *Yersinia enterocolitica*, is able to convert chorismate into salicylate, the precursor of the siderophore yersiniabactin. *J. Bacteriol.* **185**: 5648–5653.
- Pelludat, C., Hogardt, M., and Heesemann, J. (2002) Transfer of the core region genes of the *Yersinia enterocolitica* WA-C serotype O:8 high-pathogenicity island to *Y. enterocolitica*

MRS40, a strain with low levels of pathogenicity, confers a yersiniabactin biosynthesis phenotype and enhanced mouse virulence. *Infect. Immun.* **70**: 1832–1841.

Penders, J., Thijs, C., Vink, C., Stelma, F.F., Snijders, B., Kummeling, I., et al. (2006) Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* **118**: 511–521.

Perry, R.D., Balbo, P.B., Jones, H.A., Fetherston, J.D., and DeMoll, E. (1999) Yersiniabactin from *Yersinia pestis*: biochemical characterization of the siderophore and its role in iron transport and regulation. *Microbiol. Read. Engl.* **145 (Pt 5)**: 1181–1190.

Perry, R.D. and Fetherston, J.D. (2011) Yersiniabactin iron uptake: mechanisms and role in *Yersinia pestis* pathogenesis. *Microbes Infect. Inst. Pasteur* **13**: 808–817.

Pham, H.N., Ohkusu, K., Mishima, N., Noda, M., Monir Shah, M., Sun, X., et al. (2007) Phylogeny and species identification of the family Enterobacteriaceae based on dnaJ sequences. *Diagn. Microbiol. Infect. Dis.* **58**: 153–161.

Picard, B., Garcia, J.S., Gouriou, S., Duriez, P., Brahimi, N., Bingen, E., et al. (1999) The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect. Immun.* **67**: 546–553.

Pinheiro da Silva, F. (2007) CD16 promotes *Escherichia coli* sepsis through an FcR[gamma] inhibitory pathway that prevents phagocytosis and facilitates inflammation. *Nat. Med* **13**: 1368–1374.

Poisson, D.M., Borderon, J.C., Amorim-Sena, J.C., and Laugier, J. (1986) Evolution of the barrier effects against an exogenous drug-sensitive *Escherichia coli* strain after single or repeated oral administration to newborns and infants aged up to three months admitted to an intensive-care unit. *Biol Neonate* **49**: 1–7.

Poulsen, L.K. (1994) Spatial distribution of *Escherichia coli* in the mouse large intestine inferred from rRNA in situ hybridization. *Infect Immun* **62**: 5191–5194.

Priest, F.G. and Barker, M. (2010) Gram-negative bacteria associated with brewery yeasts: reclassification of *Obesumbacterium proteus* biogroup 2 as *Shimwellia pseudoproteus* gen. nov., sp. nov., and transfer of *Escherichia blattae* to *Shimwellia blattae* comb. nov. *Int. J. Syst. Evol. Microbiol.* **60**: 828–833.

Pukatzki, S., Kessin, R.H., and Mekalanos, J.J. (2002) The human pathogen *Pseudomonas aeruginosa* utilizes conserved virulence pathways to infect the social amoeba *Dictyostelium discoideum*. *Proc. Natl. Acad. Sci.* **99**: 3159–3164.

- Pukatzki, S., Ma, A.T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W.C., et al. (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 1528–1533.
- Raith, M.R., Kelty, C.A., Griffith, J.F., Schriewer, A., Wuertz, S., Mieszkina, S., et al. (2013) Comparison of PCR and quantitative real-time PCR methods for the characterization of ruminant and cattle fecal pollution sources. *Water Res.* **47**: 6921–6928.
- Rakin, A. and Heesemann, J. (1995) Yersiniabactin/pesticin receptor: a component of an iron uptake system of highly pathogenic *Yersinia*. *Contrib. Microbiol. Immunol.* **13**: 244–247.
- Rakin, A., Noeling, C., Schubert, S., and Heesemann, J. (1999) Common and specific characteristics of the high-pathogenicity island of *Yersinia enterocolitica*. *Infect. Immun.* **67**: 5265–5274.
- Rakin, A., Saken, E., Harmsen, D., and Heesemann, J. (1994) The pesticin receptor of *Yersinia enterocolitica*: a novel virulence factor with dual function. *Mol. Microbiol.* **13**: 253–263.
- Raper, K.B. and Smith, N.R. (1939) The Growth of *Dictyostelium discoideum* upon Pathogenic Bacteria. *J. Bacteriol.* **38**: 431–445.
- Rastegar Lari, A., Gold, F., Borderon, J.C., Laugier, J., and Lafont, J.P. (1990) Implantation and in vivo antagonistic effects of antibiotic-susceptible *Escherichia coli* strains administered to premature newborns. *Biol Neonate* **58**: 73–78.
- Raymond, K.N., Dertz, E.A., and Kim, S.S. (2003) Enterobactin: an archetype for microbial iron transport. *Proc. Natl. Acad. Sci. U. S. A.* **100**: 3584–3588.
- Reid, S.D., Herbelin, C.J., Bumbaugh, A.C., Selander, R.K., and Whittam, T.S. (2000) Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature* **406**: 64–67.
- Roos, V. and Klemm, P. (2006) Global gene expression profiling of the asymptomatic bacteriuria *Escherichia coli* strain 83972 in the human urinary tract. *Infect. Immun.* **74**: 3565–3575.
- Rose, R.E. (1988) The nucleotide sequence of pACYC184. *Nucleic Acids Res.* **16**: 355.
- Rupper, A. and Cardelli, J. (2001) Regulation of phagocytosis and endo-phagosomal trafficking pathways in *Dictyostelium discoideum*. *Biochim. Biophys. Acta* **1525**: 205–216.
- Russo, T.A. and Johnson, J.R. (2003) Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes Infect. Inst. Pasteur* **5**: 449–456.

- Russo, T.A., McFadden, C.D., Carlino-MacDonald, U.B., Beanan, J.M., Barnard, T.J., and Johnson, J.R. (2002) IroN functions as a siderophore receptor and is a urovirulence factor in an extraintestinal pathogenic isolate of *Escherichia coli*. *Infect. Immun.* **70**: 7156–7160.
- Savageau, M.A. (1983) *Escherichia coli* habitats, cell types, and molecular mechanisms of gene control. *Am Nat* **122**: 732–744.
- Schamberger, G.P., Phillips, R.L., Jacobs, J.L., and Diez-Gonzalez, F. (2004) Reduction of *Escherichia coli* O157:H7 populations in cattle by addition of colicin E7-producing *E. coli* to feed. *Appl Env. Microbiol* **70**: 6053–6060.
- Schierack, P., Romer, A., Jores, J., Kaspar, H., Guenther, S., Filter, M., et al. (2008) Isolation and Characterization of Intestinal *Escherichia coli* Clones from Wild Boars in Germany. *Appl. Environ. Microbiol.* **75**: 695–702.
- Schlager, T.A., Hendley, J.O., Bell, A.L., and Whittam, T.S. (2002) Clonal diversity of *Escherichia coli* colonizing stools and urinary tracts of young girls. *Infect. Immun.* **70**: 1225–1229.
- Schmidt, H. and Hensel, M. (2004) Pathogenicity islands in bacterial pathogenesis. *Clin. Microbiol. Rev.* **17**: 14–56.
- Schubert, S. (2009) Role of intraspecies recombination in the spread of pathogenicity islands within the *Escherichia coli* species. *PLoS Pathog* **5**: e1000257.
- Schubert, S., Cuenca, S., Fischer, D., and Heesemann, J. (2000) High-pathogenicity island of *Yersinia pestis* in enterobacteriaceae isolated from blood cultures and urine samples: prevalence and functional expression. *J. Infect. Dis.* **182**: 1268–1271.
- Schubert, S., Dufke, S., Sorsa, J., and Heesemann, J. (2004) A novel integrative and conjugative element (ICE) of *Escherichia coli*: the putative progenitor of the *Yersinia* high-pathogenicity island. *Mol. Microbiol.* **51**: 837–848.
- Schubert, S., Nörenberg, D., Clermont, O., Magistro, G., Wieser, A., Romann, E., et al. (2010) Prevalence and phylogenetic history of the TcpC virulence determinant in *Escherichia coli*. *Int. J. Med. Microbiol. IJMM* **300**: 429–434.
- Schubert, S., Picard, B., Gouriou, S., Heesemann, J., and Denamur, E. (2002) *Yersinia* high-pathogenicity island contributes to virulence in *Escherichia coli* causing extraintestinal infections. *Infect. Immun.* **70**: 5335–5337.
- Schubert, S., Rakin, A., Fischer, D., Sorsa, J., and Heesemann, J. (1999) Characterization of the integration site of *Yersinia* high-pathogenicity island in *Escherichia coli*. *FEMS Microbiol. Lett.* **179**: 409–414.

- Schubert, S., Rakin, A., Karch, H., Carniel, E., and Heesemann, J. (1998) Prevalence of the “high-pathogenicity island” of *Yersinia* species among *Escherichia coli* strains that are pathogenic to humans. *Infect. Immun.* **66**: 480–485.
- Sears, H.J. and Brownlee, I. (1952) Further observations on the persistence of individual strains of *Escherichia coli* in the intestinal tract of man. *J Bacteriol* **63**: 47–57.
- Sears, H.J., Brownlee, I., and Uchiyama, J.K. (1950) Persistence of individual strains of *Escherichia coli* in the intestinal tract of man. *J Bacteriol* **59**: 293–301.
- Sears, H.J., Janes, H., Saloum, R., Brownlee, I., and Lamoreaux, L.F. (1956) Persistence of individual strains of *Escherichia coli* in man and dog under varying conditions. *J Bacteriol* **71**: 370–372.
- Selander, R.K. (1986) Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Appl Env. Microbiol* **51**: 873–884.
- Skriwan, C., Fajardo, M., Hägele, S., Horn, M., Wagner, M., Michel, R., et al. (2002) Various bacterial pathogens and symbionts infect the amoeba *Dictyostelium discoideum*. *Int. J. Med. Microbiol. IJMM* **291**: 615–624.
- Skurnik, D., Bonnet, D., Bernède-Bauduin, C., Michel, R., Guette, C., Becker, J.-M., et al. (2008) Characteristics of human intestinal *Escherichia coli* with changing environments. *Environ. Microbiol.* **10**: 2132–2137.
- Skurnik, D., Lacheeb, S., Bernede, C., le Menac’h, A., Elbaz, S., Mohler, J., et al. (2009) Integrons and antibiotic resistance in phylogenetic group B2 *Escherichia coli*. *Microb. Drug Resist. Larchmt. N* **15**: 173–178.
- Slanetz, L.W. and Bartley, C.H. (1957) Numbers of enterococci in water, sewage, and feces determined by the membrane filter technique with an improved medium. *J Bacteriol* **74**: 591–595.
- Sneath, P. H., Sokal, R. R. Numerical taxonomy. San Francisco: W. H. Freeman and Co., 1973.
- Smati, M., Clermont, O., Le Gal, F., Schichmanoff, O., Jauréguy, F., Eddi, A., et al. (2013) Real-time PCR for quantitative analysis of human commensal *Escherichia coli* populations reveals a high frequency of subdominant phylogroups. *Appl. Environ. Microbiol.* **79**: 5005–5012.
- Snyder, J.A., Haugen, B.J., Buckles, E.L., Lockatell, C.V., Johnson, D.E., Sonnenberg, M.S., et al. (2004) Transcriptome of uropathogenic *Escherichia coli* during urinary tract infection. *Infect. Immun.* **72**: 6373–6381.

- Solo-Gabriele, H.M., Wolfert, M.A., Desmarais, T.R., and Palmer, C.J. (2000) Sources of *Escherichia coli* in a coastal subtropical environment. *Appl Env. Microbiol* **66**: 230–237.
- Solomon, J.M., Leung, G.S., and Isberg, R.R. (2003) Intracellular replication of *Mycobacterium marinum* within *Dictyostelium discoideum*: efficient replication in the absence of host coronin. *Infect. Immun.* **71**: 3578–3586.
- Solomon, J.M., Rupper, A., Cardelli, J.A., and Isberg, R.R. (2000) Intracellular growth of *Legionella pneumophila* in *Dictyostelium discoideum*, a system for genetic analysis of host-pathogen interactions. *Infect. Immun.* **68**: 2939–2947.
- Stamm, W.E. and Hooton, T.M. (1993) Management of urinary tract infections in adults. *N. Engl. J. Med.* **329**: 1328–1334.
- Steinert, M., Leippe, M., and Roeder, T. (2003) Surrogate hosts: protozoa and invertebrates as models for studying pathogen-host interactions. *Int. J. Med. Microbiol. IJMM* **293**: 321–332.
- Stoll, B.J., Hansen, N., Fanaroff, A.A., Wright, L.L., Carlo, W.A., Ehrenkranz, R.A., et al. (2002) Late-onset sepsis in very low birth weight neonates: the experience of the NICHD Neonatal Research Network. *Pediatrics* **110**: 285–291.
- Syed, S.A., Abrams, G.D., and Freter, R. (1970) Efficiency of various intestinal bacteria in assuming normal functions of enteric flora after association with germ-free mice. *Infect Immun* **2**: 376–386.
- Takahashi, A., Kanamaru, S., Kurazono, H., Kunishima, Y., Tsukamoto, T., Ogawa, O., and Yamamoto, S. (2006) *Escherichia coli* Isolates Associated with Uncomplicated and Complicated Cystitis and Asymptomatic Bacteriuria Possess Similar Phylogenies, Virulence Genes, and O-Serogroup Profiles. *J. Clin. Microbiol.* **44**: 4589–4592.
- Talan, D.A., Stamm, W.E., Hooton, T.M., Moran, G.J., Burke, T., Irvani, A., et al. (2000) Comparison of ciprofloxacin (7 days) and trimethoprim-sulfamethoxazole (14 days) for acute uncomplicated pyelonephritis pyelonephritis in women: a randomized trial. *JAMA* **283**: 1583–1590.
- Tenaillon, O., Skurnik, D., Picard, B., and Denamur, E. (2010) The population genetics of commensal *Escherichia coli*. *Nat. Rev. Microbiol.* **8**: 207–217.
- Torres, A.G., Redford, P., Welch, R.A., and Payne, S.M. (2001) TonB-dependent systems of uropathogenic *Escherichia coli*: aerobactin and heme transport and TonB are required for virulence in the mouse. *Infect. Immun.* **69**: 6179–6185.
- Touchon, M. (2009) Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet* **5**: e1000344.

- Tourret, J., Diard, M., Garry, L., Matic, I., and Denamur, E. (2010) Effects of single and multiple pathogenicity island deletions on uropathogenic *Escherichia coli* strain 536 intrinsic extra-intestinal virulence. *Int. J. Med. Microbiol.* **300**: 435–439.
- Unno, T., Han, D., Jang, J., Lee, S.-N., Ko, G., Choi, H.Y., et al. (2009) Absence of *Escherichia coli* Phylogenetic Group B2 Strains in Humans and Domesticated Animals from Jeonnam Province, Republic of Korea. *Appl. Environ. Microbiol.* **75**: 5659–5666.
- Urushihara, H. (2009) The cellular slime mold: eukaryotic model microorganism. *Exp. Anim. Jpn. Assoc. Lab. Anim. Sci.* **58**: 97–104.
- Valdebenito, M., Crumbliss, A.L., Winkelmann, G., and Hantke, K. (2006) Environmental factors influence the production of enterobactin, salmochelin, aerobactin, and yersiniabactin in *Escherichia coli* strain Nissle 1917. *Int. J. Med. Microbiol. IJMM* **296**: 513–520.
- Vila, J., Simon, K., Ruiz, J., Horcajada, J.P., Velasco, M., Barranco, M., et al. (2002) Are quinolone-resistant uropathogenic *Escherichia coli* less virulent? *J. Infect. Dis.* **186**: 1039–1042.
- Vimont, S., Boyd, A., Bleibtreu, A., Bens, M., Goujon, J.-M., Garry, L., et al. (2012) The CTX-M-15-producing *Escherichia coli* clone O25b: H4-ST131 has high intestine colonization and urinary tract infection abilities. *PloS One* **7**: e46547.
- Visca, P., Filetici, E., Anastoasio, M.P., Vetriani, C., Fantasia, M., and Orsi, N. (1991) Siderophore production by *Salmonella* species isolated from different sources. *FEMS Microbiol. Lett.* **63**: 225–231.
- Vollaard, E.J. and Clasener, H.A. (1994) Colonization resistance. *Antimicrob Agents Chemother* **38**: 409–414.
- Vollmerhausen, T.L., Ramos, N.L., Gündogdu, A., Robinson, W., Brauner, A., and Katouli, M. (2011) Population structure and uropathogenic virulence-associated genes of faecal *Escherichia coli* from healthy young and elderly adults. *J. Med. Microbiol.* **60**: 574–581.
- Walk, S.T., Alm, E.W., Gordon, D.M., Ram, J.L., Toranzos, G.A., Tiedje, J.M., and Whittam, T.S. (2009) Cryptic lineages of the genus *Escherichia*. *Appl. Environ. Microbiol.* **75**: 6534–6544.
- Walk, S.T., Mladonicky, J.M., Middleton, J.A., Heidt, A.J., Cunningham, J.R., Bartlett, P., et al. (2007) Influence of Antibiotic Selection on Genetic Composition of *Escherichia coli* Populations from Conventional and Organic Dairy Farms. *Appl. Environ. Microbiol.* **73**: 5982–5989.
- Wandersman, C. and Delepelaire, P. (2004) Bacterial iron sources: from siderophores to hemophores. *Annu. Rev. Microbiol.* **58**: 611–647.

- Wang, R.F. and Kushner, S.R. (1991) Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* **100**: 195–199.
- Ward, P.P. and Conneely, O.M. (2004) Lactoferrin: role in iron homeostasis and host defense against microbial infection. *Biometals Int. J. Role Met. Ions Biol. Biochem. Med.* **17**: 203–208.
- Watt, S. (2003) *Escherichia coli* strains from pregnant women and neonates: intraspecies genetic distribution and prevalence of virulence factors. *J Clin Microbiol* **41**: 1929–1935.
- Welch, R.A., Burland, V., Plunkett, G., Redford, P., Roesch, P., Rasko, D., et al. (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **99**: 17020–17024.
- Whittam, T.S., Ochman, H., and Selander, R.K. (1983) Multilocus genetic structure in natural populations of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **80**: 1751–1755.
- Wieser, A., Romann, E., Magistro, G., Hoffmann, C., Nörenberg, D., Weinert, K., and Schubert, S. (2010) A multiepitope subunit vaccine conveys protection against extraintestinal pathogenic *Escherichia coli* in mice. *Infect. Immun.* **78**: 3432–3442.
- Wiles, T.J., Kulesus, R.R., and Mulvey, M.A. (2008) Origins and virulence mechanisms of uropathogenic *Escherichia coli*. *Exp. Mol. Pathol.* **85**: 11–19.
- Wirth, T., Falush, D., Lan, R., Colles, F., Mensa, P., Wieler, L.H., et al. (2006) Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol. Microbiol.* **60**: 1136–1151.
- Zhang, L., Foxman, B., and Marrs, C. (2002) Both urinary and rectal *Escherichia coli* isolates are dominated by strains of phylogenetic group B2. *J Clin Microbiol* **40**: 3951–3955.

Place de la structure génétique de l'espèce *Escherichia coli* dans l'état de son commensalisme intestinal et dans l'expression de sa virulence

Résumé

Escherichia coli est le commensal aérobie le plus fréquent du tube digestif de l'homme et des animaux à sang chaud et le bacille à Gram négatif pathogène opportuniste le plus souvent impliqué dans les infections intestinales et extra intestinales de l'homme. C'est une espèce clonale chez laquelle 4 groupes phylogénétiques principaux, A, B1, B2 et D ont été décrits. L'objectif de cette thèse est d'étudier l'adaptation de *E. coli* et les rapports de cette adaptation avec la structure génétique de l'espèce caractérisée par les 4 groupes phylogénétiques dans deux circonstances : le commensalisme intestinal de l'homme et de plusieurs espèces animales sauvages et domestiques, herbivores et omnivores d'une part et la virulence extra-intestinale mesurée par l'expression des gènes codants pour un sidérophore, la yersiniabactine, dont les gènes sont situés au sein de l'îlot de pathogénicité HPI (PAI IV).

La répartition dans les 4 groupes phylogénétiques des souches commensales du tube digestif de 100 hommes et de 137 animaux a été étudiée par une technique de PCR en temps réel originale. Trois principaux entérocolotypes, correspondant à des associations préférentielles de phylogroupes ont été ainsi décrits comme plus fréquents en fonction de la nature des hôtes. Chez l'homme, les souches du groupe B2 ont été retrouvées exclusives chez 15 % des individus et ont été clairement distinctes des souches B2 des animaux sauvages par la plus grande fréquence de leurs facteurs de virulences (*sfa/foc* et *pks*).

L'effet du fond génétique des sous groupes II, III et IX du groupe B2 sur l'expression de la virulence liée au HPI a été étudié dans un modèle murin de virulence extra-intestinale et dans un modèle d'amibe sociale *Dictyostelium discoideum*, pouvant être assimilé à un macrophage. Le HPI chez *E. coli* interagit avec la clonalité de l'espèce qui s'exprime par l'existence des sous-groupes de B2. Dans les modèles de virulence que nous avons développés, les mêmes gènes ont, en fonction du fond génétique des différents isolats naturels, des effets différents.

Mots clés

Escherichia coli, HPI, phylogroupes, commensalisme, PCR quantitative, entérocolotype, structures des populations commensales.